(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 February 2008 (21.02.2008)

PCT

(10) International Publication Number
WO 2008/020335 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/IB2007/003693

(22) International Filing Date: 11 June 2007 (11.06.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/812145 9 June 2006 (09.06.2006) US

(71) Applicant *(for all designated States except US)*: **NOVARTIS AG** [CH/CH]; Lichtstrasse 35, CH-4056 Basel (CH).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MAIONE, Domenico** [IT/IT]; C/o Novartis Vaccines And Diagnostics Inc., P.o. Box 8097, Emeryville, CA 94662-8097 (US). **NORAIS, Nathalie** [IT/IT]; C/o Novartis Vaccines And Diagnostics Inc., P.o. Box 8097, Emeryville, CA 94662-8097 (US). **GRANDI, Guido** [IT/IT]; C/o Novartis Vaccines And Diagnostics Inc., P.o. Box 8097, Emeryville, CA 94662-8097 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *without international search report and to be republished upon receipt of that report*

WO 2008/020335 A2

(54) Title: IMMUNOGENIC COMPOSITIONS FOR STREPTOCOCCUS AGALACTIAE

(57) Abstract: The invention relates to immunogenic polypeptides derived from epitopes in a Streptococcus agalactiae ("GBS") protein GBS 80 and their use as prophylactic, diagnostic and therapeutic compositions. The invention also provides nucleic acids encoding the immunogenic polypeptides. Also provided are vectors useful for making such immunogenic polypeptides and host cells transformed with such vectors. In particular, the invention relates to a group immunogenic polypeptides derived from GBS 80. The compositions may include one or more of the immunogenic polypeptides either alone or with other antigenic components. For example, the immunogenic polypeptides may be combined with other GBS antigens to provide therapeutic compositions with broader range. In addition, the immunogenic polypeptides may also include flanking portions of the GBS 80 protein.

# IMMUNOGENIC COMPOSITIONS FOR STREPTOCOCCUS AGALACTIAE

## FIELD OF THE INVENTION

**[0001]** The invention relates to immunogenic polypeptides derived from a *Streptococcus agalactiae* ("GBS") protein GBS 80 and their use as diagnostic, prophylactic, and therapeutic compositions. In particular, the invention relates to a group of immunogenic polypeptides derived from GBS 80. The compositions may include one or more of the immunogenic polypeptides either alone or with other immunogenic components. For example, the immunogenic polypeptides may be combined with other GBS antigens to provide therapeutic compositions with broader range. In addition, the immunogenic polypeptides may also include flanking portions of the GBS 80 protein.

## BACKGROUND OF THE INVENTION

**[0002]** GBS has emerged in the last 20 years as the major cause of neonatal sepsis and meningitis that affect 0.5-3 per 1000 live births, and an important cause of morbidity among the older age group affecting 5-8 per 100,000 of the population. Current disease management strategies rely on intrapartum antibiotics and neonatal monitoring which have reduced neonatal case mortality from >50% in the 1970's to less than 10% in the 1990's. Nevertheless, there is still considerable morbidity and mortality and the management is expensive. 15-35% of pregnant women are asymptomatic carriers and at high risk of transmitting the disease to their babies. Risk of neonatal infection is associated with low serotype specific maternal antibodies and high titers are believed to be protective. In addition, invasive GBS disease is increasingly recognized in elderly adults with underlying disease such as diabetes and cancer.

**[0003]** The "B" in "GBS" refers to the Lancefield classification, which is based on the antigenicity of a carbohydrate which is soluble in dilute acid and called the C carbohydrate. Lancefield identified 13 types of C carbohydrate, designated A to O, that could be serologically differentiated; the organisms that most commonly infect humans are found in groups A, B, D, and G. Within group B, strains can be divided into at least 9 serotypes (Ia, Ib, Ia/c, II, III, IV, V, VI, VII and VIII) based on the structure of their polysaccharide capsule. In the past, serotypes Ia, Ib, II, and III were equally prevalent in normal vaginal carriage and early onset sepsis in newborns. Type V GBS has emerged as an important cause

of GBS infection in the USA, however, and strains of types VI and VIII have become prevalent among Japanese women.

**[0004]** The genome sequence of a serotype V strain 2603 V/R has been published (Ref. 1) and various polypeptides for use a vaccine antigens have been identified (Ref. 2). The vaccines currently in clinical trials, however, are based on polysaccharide antigens. These suffer from serotype specificity and poor immunogenicity, and so there is a need for effective vaccines against *S. agalactiae* infection.

**[0005]** It is an object of the invention to provide improved compositions for providing immunity against, and treatment of, GBS disease and/or infection. The compositions are based on a group of immunogenic polypeptides derived from GBS 80.

## SUMMARY OF THE INVENTION

**[0006]** Applicants have discovered that an immunogenic GBS antigen, GBS 80, is particularly suitable for immunization purposes, which may be used in combination with other GBS antigens. Applicants have identified four regions within GBS 80 that are of particular interest given their demonstrated antigenic qualities.

**[0007]** One aspect of the present invention provides an immunogenic composition comprising an immunogenic polypeptide from GBS 80 or a fragment thereof, wherein said immunogenic polypeptide is a fragment of GBS 80 that includes one of the regions identified in this application (especially SEQ ID NO:7-12 or antigenic fragment thereof) and may include additional portions of GBS 80. The length of the fragment may vary depending on the amino acid sequence of the specific immunogenic polypeptide, but the fragment is preferably at least 7 consecutive amino acids, (*e.g.* 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 or more).

**[0008]** The immunogenic polypeptides may include polypeptide sequences having sequence identity to the identified immunogenic polypeptides (especially SEQ ID NO:7-12 or antigenic fragments thereof). The degree of sequence identity may vary depending on the amino acid sequence in question, but is preferably greater than 50% (*e.g.* 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or more). Polypeptides having sequence identity include homologs, orthologs, allelic variants and functional mutants of the identified GBS 80 immunogenic polypeptides. The immunogenic

polypeptides may include polypeptide sequences encoded by nucleic acid sequences that hybridize under high stringency wash conditions (see below for representative conditions) to nucleic acids encoding an identified immunogenic polypeptide (especially SEQ ID NO: 7-12 or antigenic fragments thereof).

**[0009]** With regard to the immunogenic polypeptide of SEQ ID NO:7, polypeptides of particular interest include polypeptides having a region of limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 270 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:7. In preferred embodiments, the contiguous region will be "over 268," "over 266," "over 264," "over 262," "over 260," "over 250," "over 240," "over 230," "over 220," "over 210," "over 200," "over 180," "over 160," "over 140," "over 120," "over 100," "over 80," "over 60," "over 50," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14," "over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous identity is desired and the phrase "or fewer" may be replaced with "to 200," "to 180," "to 160," "to 140," "to 120," "to 100," "to 80," "to 60," "to 50," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0010]** With regard to the immunogenic polypeptide of SEQ ID NO:8, polypeptides of particular interest include polypeptides having a region of limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 270 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:8. In preferred embodiments, the contiguous region will be "over 268," "over 266," "over 264," "over 262," "over 260," "over 250," "over 240," "over 230," "over 220," "over 210," "over 200," "over 180," "over 160," "over 140," "over 120," "over 100," "over 80," "over 60," "over 50," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14," "over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous

identity is desired and the phrase "or fewer" may be replaced with "to 200," "to 180," "to 160," "to 140," "to 120," "to 100," "to 80," "to 60," "to 50," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0011]** With regard to the immunogenic polypeptide of SEQ ID NO:9, polypeptides of particular interest include polypeptides having a region of limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 211 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:9. In preferred embodiments, the contiguous region will be "over 208," "over 206," "over 204," "over 202," "over 200," "over 190," "over 180," "over 170," "over 160," "over 150," "over 140," "over 130," "over 120," "over 100," "over 80," "over 74," "over 72," "over 70," "over 65," "over 60," "over 55," "over 50," "over 45," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14," "over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous identity is desired and the phrase "or fewer" may be replaced with "to 147," "to 146," "to 145," "to 140," "to 130," "to 120," "to 100," "to 80," "to 74," "to 72," "to 70," "to 65," "to 60," "to 55," "to 50," "to 45," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0012]** With regard to the immunogenic polypeptide of SEQ ID NO:10, polypeptides of particular interest include polypeptides having a region of limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 76 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:10. In preferred embodiments, the contiguous region will be "over 74," "over 72," "over 70," "over 65," "over 60," "over 55," "over 50," "over 45," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14,"

"over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous identity is desired and the phrase "or fewer" may be replaced with "to 70," "to 65," "to 60," "to 55," "to 50," "to 45," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0013]** With regard to the immunogenic polypeptide of SEQ ID NO:11, polypeptides of particular interest include polypeptides having a region of limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 270 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:11. In preferred embodiments, the contiguous region will be "over 268," "over 266," "over 264," "over 262," "over 260," "over 250," "over 240," "over 230," "over 220," "over 210," "over 200," "over 180," "over 160," "over 140," "over 120," "over 100," "over 80," "over 60," "over 50," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14," "over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous identity is desired and the phrase "or fewer" may be replaced with "to 200," "to 180," "to 160," "to 140," "to 120," "to 100," "to 80," "to 60," "to 50," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0014]** With regard to the immunogenic polypeptide of SEQ ID NO:12, polypeptides of particular interest include polypeptides having a region of limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 270 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:12. In preferred embodiments, the contiguous region will be "over 268," "over 266," "over 264," "over 262," "over 260," "over 250," "over 240," "over 230," "over 220," "over 210," "over 200," "over 180," "over 160," "over 140," "over 120,"

"over 100," "over 80," "over 60," "over 50," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14," "over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous identity is desired and the phrase "or fewer" may be replaced with "to 200," "to 180," "to 160," "to 140," "to 120," "to 100," "to 80," "to 60," "to 50," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0015]** With regard to the immunogenic polypeptide of SEQ ID NO:9, additional polypeptides of particular interest include polypeptides having a region of limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) to SEQ ID NO:2, wherein the region includes SEQ ID NO:9 and extends no more than 47 amino acids upstream of SEQ ID NO:9. In preferred embodiments, the region of contiguous alignment will extend no more than 46, 44, 42, 40, 35, 30, 25, 20, 15, 10, 8, 7, 5, 4, 3, 2, or 1 amino acid(s) upstream. In some embodiments, the region of contiguous alignment will begin with the N-terminal end of SEQ ID NO:9.

**[0016]** With regard to the immunogenic polypeptide of SEQ ID NO:10, additional polypeptides of particular interest include polypeptides having a region of limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) to SEQ ID NO:2, wherein the region includes SEQ ID NO:10 and extends no more than 56 amino acids upstream. In preferred embodiments, the region of contiguous alignment will extend no more than 54, 52, 50, 48, 46, 44, 42, 40, 35, 30, 25, 20, 15, 10, 8, 7, 5, 4, 3, 2, or 1 amino acid(s) upstream. In some embodiments, the region of contiguous alignment will begin with the N-terminal end of SEQ ID NO:10.

**[0017]** In preferred embodiments, the polypeptides of the present invention will be capable of generating an immune response in a target organism such as a bird or a mammal, preferably a human subject. More preferably, the polypeptides will provide a target organism passive immunity and/or active immunity.

**[0018]** Additional embodiments of the polypeptides of the present invention may be found throughout the specification. By way of example, the polypeptides may further comprise targeting sequences such as secretion sequences, purification sequences, and fusion proteins including without limitation other immunogenic polypeptides, proteins that improve stability of the polypeptide, retention of the polypeptide within the subject, or antigenicity of the polypeptide.

**[0019]** In some embodiments, the polypeptide compositions of the present invention may additionally include other immunogenic polypeptides from GBS 80 (including without limitation polypeptides and polysaccharides) or other pathogens.

**[0020]** As described more fully below, additional aspects of the present invention include methods of using the foregoing polypeptides as (a) medicaments for treating or preventing infection due to *Streptococcus* bacteria; (b) diagnostics or immunodiagnostic assays for detecting the presence of *Streptococcus* bacteria or of antibodies raised against *Streptococcus* bacteria; and/or (c) reagents which can raise antibodies against *Streptococcus bacteria.*

**[0021]** Another aspect of the present invention includes methods of screening and/or testing peptides of the present invention for generation of an immune response, active immunization or passive immunization in a target organism. In some embodiments, the invention will involve contacting or administering the polypeptide composition of the present invention to the target organism and detecting antibodies in the target organism that recognize the polypeptide composition. In preferred embodiments, the target organism will be challenged with a *Streptococcus* bacterium to determine whether the target organism has active immunity or passive immunity. Such methods of screening may be applied to any of the compositions of the present invention including, without limitation, immunogenic polypeptides and pharmaceutical compositions for immunogenicity or antigenicity. A preferred embodiment of such screening methods includes providing an immunogenic polypeptide and screening the polypeptide for antigenicity or immunogenicity. Where more than one immunogenic polypeptide is to be screened, a criterion may be applied to select one or more immunogenic polypeptides for further use. Such criteria may be used to select among two or more immunogenic polypeptides, three or more immunogenic polypeptides, five or more immunogenic polypeptides, ten or more immunogenic polypeptides, or twenty or more immunogenic polypeptides.

**[0022]** Another aspect of the present invention is nucleic acids encoding any of the polypeptides of the present invention. In certain embodiments, such nucleic acids may be in an isolated or in recombinant form. In some embodiments, the nucleic acids encoding any of the foregoing polypeptides may be in a vector. In some embodiments, such nucleic acids may be operably linked to a promoter which preferably is operable in the host organism in which the polypeptide is to be expressed. In various embodiments, the promoter may be a constitutive promoter, a regulatable promoter, or an inducible promoter. Additional embodiments are described more fully below regarding expression vectors including nucleic acids of the present invention.

**[0023]** Another aspect of the present invention provides pharmaceutical compositions that include the polypeptides, antibodies, or nucleic acids of the present invention in a therapeutically effective amount (or an immunologically effective amount in a vaccine). In certain embodiments, the pharmaceutical compositions will be vaccines. The pharmaceutical vaccines may also have pharmaceutically acceptable carriers including adjuvants.

**[0024]** Additional aspects and embodiments may be found throughout the specification. The specification is not intended as a limitation of the scope of the present invention, but rather as examples of the aspects and embodiments of the present invention. One of skill in the art can infer additional embodiments from the description provided.

## BRIEF DESCRIPTION THE FIGURES

**[0025]** Figure 1 shows the predicted fragments from the recombinantly produced GBS 80. The recombinantly produced protein has the N-terminal leader peptide removed (37 amino acids) and the C-terminal cell wall anchor and transmembrane region removed.

**[0026]** Figure 2 shows the predicted mass-to-charge ratio for each of the predicted fragments identified in Figure 1.

**[0027]** Figure 3 shows a summary of western blot and FACs analysis conducted with six monoclonal antibodies directed to GBS 80 used to identify the immunogenic polypeptides herein.

**[0028]** Figure 4 shows FACs analysis graphs of the six monoclonal antibodies and a polyclonal antibody serum.

[0029] Figure 5 shows the general scheme used to identify fragments produced in partial digests of recombinantly produced GBS 80.

[0030] Figure 6 shows western blots of partial Asp-N digests of recombinantly produced GBS. On the left is a western blot using the 9A4/77 monoclonal antibody and on the right is a western blot using the M3/88 monoclonal antibody.

[0031] Figure 7 shows a Coomassie Blue stained SDS-PAGE of partial digests recombinantly produced GBS 80 using two different proteases, Asp-N and Arg-C. GBS 80 F and GBS 80 3 correspond to two different conformations of GBS 80 which have different protease sensitivities. The lanes are as labeled on the figure.

[0032] Figure 8 shows a pair of western blots of the two conformers of GBS 80 partially digested with either Asp-N or Arg-C. On the left is a western blot using the 9A4/77 monoclonal antibody and on the right is a western blot using the M3/88 monoclonal antibody. The lanes are as labeled on the figure.

[0033] Figure 9 shows an SDS-PAGE of the partial digests of boiled samples of GBS 80: 1) an Arg-C partial digest of GBS 80 3, 2) an Arg-C partial digest of GBS 80 F, 3) an Asp-N partial digest of GBS 80 F, and 4) GBS 80 F (no digest). M indicates lanes with protein markers of the sizes indicated along the left of the gel image.

[0034] Figure 10 shows the results of the western blot epitope mapping of the 9A4/77 monoclonal antibody and the M3/88 monoclonal antibody. The full length GBS 80 protein is shown schematically along the top with numbers indicating the amino acid position. Each protein fragment identified by MALDI-TOF from Figure 9 is show below the full length GBS 80 protein with the corresponding fragment number. Along the left are two columns indicating which of the fragments were observed in the western blots with the two antibodies – N is 9A4/77 and C is M3/88. The two circles indicate the regions bound by each antibody.

[0035] Figure 11 shows the sequence of the recombinantly produced GBS 80 protein (note: the recombinant GBS 80 has had the N-terminal leader sequence removed and replaced with a leading methionine residue, so amino acid 1 corresponds to 37 in the full length GBS 80 protein). Three immunogenic polypeptides are shown in the figure. The yellow region is recognized by 9A4/77. The cyan and green region is bound by M3/88 and the green region is the core region bound by M3/88.

**[0036]** Figure 12 shows four western blots of the two conformers of GBS 80 partially digested with either Asp-N or Arg-C. The antibodies use to generate the western blots are indicated above the respective western blot. The lanes are as labeled on the figure.

**[0037]** Figure 13 shows the sequence of the recombinantly produced GBS 80 protein (again note: the recombinant GBS 80 has had the N-terminal leader sequence removed and replaced with a leading methionine residue, so amino acid 1 corresponds to 37 in the full length GBS 80 protein). Three immunogenic polypeptides are shown in the figure. The yellow region is recognized by 19G4/78 and 19F6/77. The cyan and green region are bound by M1/77 and M2/77 while the green region represents the core region bound by the two antibodies.

**[0038]** Figure 14 shows a schematic of the layout of a peptide microarray used to further identify immunogenic polypeptides in GBS 80. The control peptides are around the edges of the chip labeled with roman numerals I-VI. The GBS 80 peptides are numbered 1-80 as set out in Table 3 below (note: there are no peptides in positions 28-36).

**[0039]** Figure 15 shows three peptide microarrays on a slide after fluorescent labeling. Control peptides are indicated with dashed circles and GBS 80 peptides are indicated with solid circles. Peptides number 73 and 75 were both bound by the monoclonal antibody 9A4/77 in all three arrays on the slide.

**[0040]** Figure 16 shows the sequence of the recombinantly produced GBS 80 protein. Three immunogenic polypeptides are shown in the figure. The immunogenic polypeptide identified from the microarray epitope mapping is shown as cyan highlighting.

## DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

**[0041]** An aspect of the present invention provides fragments and sub fragments of the proteins and protein fragments disclosed in international patent applications WO04/041157 and WO05/028618 (the "International Applications"), wherein the fragments comprise at least one immunogenic polypeptide.

**[0042]** Thus, if the length of any particular protein or protein fragment sequence disclosed in the International Applications is x amino acids, the present invention provides fragments of at most x-1 amino acids of that protein. The fragment may be shorter than this (e.g., x-2, x-3, x-4,...), and is preferably 100 amino acids or less (e.g., 90 amino acids, 80 amino acids etc.).

The fragment may be as short as 3 amino acids, but is preferably longer (e.g., up to 5, 6, 7, 8, 9, 10, 12, 15, 20, 25, 30, 35, 40, 50, 75, or 100 amino acids).

**[0043]** Preferred fragments comprise the GBS 80 immunogenic polypeptides disclosed below, or sub-sequences thereof. The fragments may be longer than those disclosed below e.g., where a fragment runs from amino acid residue p to residue q of a protein, the invention also relates to fragments from residue (p-1), (p-2), or (p-3) to residue (q+l), (q+2), or (q+3), up to 1 amino acid less that the fragments disclosed in the International Applications.

**[0044]** The invention also provides polypeptides that are homologous (i.e., have sequence identity) to these fragments. Depending on the particular fragment, the degree of sequence identity is preferably greater than 50% (e.g., 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5% or more). These homologous polypeptides include mutants and allelic variants of the fragments. Identity between the two sequences is preferably determined by the Smith- Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), preferably using an affine gap search with parameters gap open penalty=12 and gap extension penalty=1.

**[0045]** The invention also provides proteins comprising one or more of the above-defined fragments.

**[0046]** The invention is subject to the proviso that it does not include within its scope proteins limited to any of the full length protein or protein fragment sequences disclosed in the International Applications (i.e., SEQ ID NOs: 1 and 2 of WO04/041157 and SEQ ID NOs: 1-9 of WO05/028618).

**[0047]** The proteins of the invention can, of course, be prepared by various means (e.g., recombinant expression, purification from cell culture, chemical synthesis etc.) and in various forms (e.g., native, C- terminal and/or N-terminal fusions etc.). They are preferably prepared in substantially pure form (i.e., substantially free from other GBS or host cell proteins, with the understanding that they may later be combined with antigens from GBS or other pathogens to create combination vaccines). Short polypeptides are preferably produced using chemical peptide synthesis.

**[0048]** According to a further aspect, the invention provides antibodies which recognize the fragments of the invention, with the proviso that the invention does not include within its

scope antibodies which recognize any of the complete protein sequences in the International Applications. The antibodies may be polyclonal or monoclonal, and may be produced by any suitable means. Example 2 provides examples of monoclonal and polyclonal antibodies that recognize certain immunogenic polypeptides of the present invention.

**[0049]** The invention also provides proteins comprising peptide sequences recognized by these antibodies. These peptide sequences will, of course, include fragments of the GBS 80 protein and protein fragments in the International Applications, but will also include peptides that mimic the antigenic structure of the GBS 80 peptides when bound to immunoglobulin.

**[0050]** According to a further aspect, the invention provides nucleic acids encoding the fragments and proteins of the invention, with the proviso that the invention does not include within its scope nucleic acid encoding any of the full length protein or protein fragment sequences in the International Applications. The nucleic acids may be as short as 10 nucleotides, but are preferably longer (e.g., up to 10, 12, 15, 18, 20, 25, 30, 35, 40, 50, 75, or 100 nucleotides).

**[0051]** In addition, the invention provides nucleic acid comprising sequences homologous (i.e., having sequence identity) to these sequences. The degree of sequence identity is preferably greater than 50% (e.g., 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.5% or more). Furthermore, the invention provides nucleic acid which can hybridize to these sequences, preferably under "high stringency" conditions (e. g., at least one wash at 65° C in a 0. 1 x SSC, 0.5 % SDS for 15 minutes).

**[0052]** It should also be appreciated that the invention provides nucleic acid comprising sequences complementary to those described above (e.g., for antisense or probing purposes).

**[0053]** Nucleic acids according to the invention can, of course, be prepared in many ways (e.g., by chemical synthesis, from genomic or cDNA libraries, from the organism itself etc.) and can take various forms (e.g., single stranded, double stranded, vectors, probes etc.). In addition, the term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also peptide nucleic acids (PNA), etc. According to a further aspect, the invention provides vectors comprising nucleotide sequences of the invention (e.g., expression vectors) and host cells transformed with such vectors.

**[0054]** According to a further aspect, the invention provides compositions comprising protein, antibody, and/or nucleic acid according to the invention. These compositions may be suitable as vaccines, for instance, or as other prophylactic agents, or as diagnostic reagents, or as immunogenic compositions. Therefore, another aspect of the present invention includes the use of nucleic acid, protein, or antibody according to the invention in the manufacture of: (i) a medicament for treating or preventing infection due to *Streptococcus* bacteria; (ii) a diagnostic reagent for detecting the presence of *Streptococcus* bacteria or of antibodies raised against *Streptococcus* bacteria; and/or (iii) a reagent which can raise antibodies against *Streptococcus* bacteria. Said *Streptococcus* bacteria may be any species or strain (such as *Streptococcus pyogenes* and *S. pneumonia*) but are preferably the Lancefield-streptococci strains, more preferably the Lancefield group B strains and most preferably *Streptococcus agalactiae*, in each of the foregoing, the bacteria are limited to those having a GBS-type pilus and therefore a GBS 80 homolog. The invention also provides a method of treating a patient, comprising administering to the patient a therapeutically effective amount of nucleic acid, protein, and/or antibody according to the invention, According to further aspects, the invention provides various processes, for example:

**[0055]** A process for producing proteins of the invention is provided, comprising the step of culturing a host cell according to the invention under conditions which induce protein expression. A process for producing protein or nucleic acid of the invention is provided, wherein the protein or nucleic acid is synthesized in part or in whole using chemical means. A process for detecting polynucleotides of the invention is provided, comprising the steps of: (a) contacting a nucleic probe according to the invention with a biological sample under hybridizing conditions to form duplexes; and (b) detecting said duplexes. In preferred examples, the detection of the duplex involves amplification of the nucleic acid detected, more preferably through RT-PCR. A process for detecting proteins of the invention is provided, comprising the steps of: (a) contacting an antibody according to the invention with a biological sample under conditions suitable for the formation of an antibody-antigen complexes; and (b) detecting said complexes.

**[0056]** A summary of standard techniques and procedures which may be employed in order to perform the invention (e.g., to utilize the disclosed sequences for vaccination or diagnostic purposes) follows. This summary is not a limitation on the invention but, rather, gives examples which may be used, but which are not required.

General

[0057] The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, *DNA Cloning*, Volumes I and II (D.N Glover ed. 1985); *Oligonucleotide Synthesis* (MT Gait ed, 1984); *Nucleic Acid Hybridization* (B.D. Hames & ST Higgins eds. 1984); *Transcription and Translation* (B.D. Hames & ST Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); *Immobilized Cells and Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide to Molecular Cloning* (1984); the Methods in Enzymology series (Academic Press, Inc.), especially volumes 154 & 155; *Gene Transfer Vectors for Mammalian Cells* (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Mayer and Walker, eds. (1987), *Immunochemical Methods in Cell and Molecular Biology* (Academic Press, London); Scopes, (1987) *Protein Purification: Principles and Practice*, Second Edition (Springer-Verlag, N.Y.), *Handbook of Experimental Immunology*, Volumes I-IV (D.M. Weir and C. C. Blackwell eds 1986), *Remington's Pharmaceutical Sciences*, Mack Publishing Company, Easton, Pa., 19th Edition (1995); *Methods In Enzymology* (S. Colowick and N. Kaplan, eds., Academic Press, Inc.); and *Handbook of Experimental Immunology*, Vols. I-IV (D.M. Weir and C.C. Blackwell, eds., 1986, Blackwell Scientific Publications); Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); *Handbook of Surface and Colloidal Chemistry* (Birdi, K.S. ed., CRC Press, 1997); *Short Protocols in Molecular Biology*, 4th ed. (Ausubel *et al.* eds., 1999, John Wiley & Sons); *Molecular Biology Techniques: An Intensive Laboratory Course*, (Ream *et al.*, eds., 1998, Academic Press); *PCR* (*Introduction to Biotechniques Series*), 2nd ed. (Newton & Graham eds., 1997, Springer Verlag); and Peters and Dalrymple, *Fields Virology* (2d ed), Fields *et al.* (eds.), B.N. Raven Press, New York, NY.

[0058] Standard abbreviations for nucleotides and amino acids are used in this specification.

[0059] All publications, patents, and patent applications cited herein are incorporated in full by reference.

Definitions

**[0060]** A composition containing X is "substantially free of" Y when at least 85% by weight of the total X+Y in the composition is X. Preferably, X comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95% or even 99% by weight.

**[0061]** The term "comprising" means "including" as well as "consisting'" e. g., a composition "comprising" X may consist exclusively of X or may include something additional to X, such as X+Y.

**[0062]** The term "antigenic determinant" includes B-cell epitopes and T- cell epitopes.

**[0063]** The term "heterologous" refers to two biological components that are not found together in nature. The components may be host cells, genes, or regulatory regions, such as promoters. Although the heterologous components are not found together in nature, they can function together, as when a promoter heterologous to a gene is operably linked to the gene. Another example is where a meningococcal sequence is heterologous to a mouse host cell. A further example would be two epitopes from the same or different proteins which have been assembled in a single protein in an arrangement not found in nature.

**[0064]** An "origin of replication" is a polynucleotide sequence that initiates and regulates replication of polynucleotides, such as an expression vector. The origin of replication behaves as an autonomous unit of polynucleotide replication within a cell, capable of replication under its own control. An origin of replication may be needed for a vector to replicate in a particular host cell. With certain origins of replication, an expression vector can be reproduced at a high copy number in the presence of the appropriate proteins within the cell. Examples of origins are the autonomously replicating sequences, which are effective in yeast; and the viral T-antigen, effective in COS-7 cells.

**[0065]** The term "a polypeptide having a region of limited, contiguous sequence identity of at least X percent over Y {or fewer} amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:Z" as used herein means that the polypeptide has a percent identity of at least X percent (e.g., at least 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5%) when compared to SEQ ID NO:2, and the region of alignment in SEQ ID NO:2 includes SEQ ID NO:Z and is limited and contiguous. In the context of this phrase,

contiguous means that when the polypeptide's sequence is aligned with SEQ ID NO:2 there are no gaps in the alignment or if there are, the amino acids across the gap are considered non-identical amino acids for the purpose of calculating the percent identity. In the context of this phrase, limited means that the polypeptide may be longer than Y amino acids, but that the polypeptide when aligned to the sequence of GBS 80 will have not have a region of alignment that is longer than Y amino acids. For the avoidance of doubt, where the region of alignment is flanked by amino acids that are not conserved, they are not included in the calculation of the length Y even if they could be included and still meet the percent identity. For example, the term "a polypeptide having a region of limited contiguous sequence identity of at least 90 percent over 100 amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:7" would include a polypeptide that has a contiguous region of 100 amino acids that has 97% identity to GBS 80 (SEQ ID NO:2) and includes SEQ ID NO:7 even when the polypeptide is longer than 100 amino acids as long as the flanking amino acids are not conserved even though the flanking amino acids could be included and still be at least 90 percent identical (i.e., a pair of sequences that are 102 amino acids in length and have 97 conserved amino acids would have a 95% identity). Thus, the this term would include polypeptides that have additional sequences fused to the immunogenic polypeptide such as signal peptides, additional epitopes (including other epitopes from GBS 80 as long as they are not contiguous with the immunogenic polypeptide or are within the contiguous region), and other proteins and polypeptides that one of skill in the art may desire.

**[0066]** The term "a polypeptide having a region of limited, contiguous sequence identity of at least X percent to SEQ ID NO:2, wherein the region includes SEQ ID NO:Y and extends no more than Z amino acids {upstream/downstream} of SEQ ID NO:Y" as used herein means that the polypeptide has a region that is at least X percent identical (e.g., at least 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5%) when compared to SEQ ID NO:2, and the region of alignment in SEQ ID NO:2 includes SEQ ID NO:Y and is limited and contiguous. In the context of this phrase, contiguous means that when the polypeptide's sequence is aligned with SEQ ID NO:Y there are no gaps in the alignment or if there are, the amino acids across the gap are considered non-identical amino acids for the purpose of calculating the percent identity. In the context of this phrase, limited means that the polypeptide may extend upstream (i.e., N-terminal to SEQ ID NO:Y) or downstream (i.e., C-terminal to SEQ ID NO:Y) longer than Z amino acids, but that the polypeptide when aligned to the sequence of GBS 80 will have not have a region of alignment upstream or downstream,

respectively, that extends more than Z amino acids from SEQ ID NO:Y. For the avoidance of doubt, where the region of alignment is flanked by amino acids that are not conserved, they are not included in the calculation of the length Z even if they could be included and still meet the percent identity. For example, the term "a polypeptide having a region of limited, contiguous sequence identity of at least 90 percent to SEQ ID NO:2 wherein the region includes SEQ ID NO:7 and extends no more than 50 amino acids upstream of SEQ IS NO:7" would include a polypeptide that has a contiguous region of 50 amino acids immediately upstream of the region that is identical to SEQ ID NO:7 that has 97% identity to GBS 80 and even when the polypeptide extends upstream further than 50 amino acids as long as the amino acids immediately upstream of the 50 amino acid stretch are not conserved even though the flanking amino acids could be included and still be at least 90 percent identical.

GBS 80

**[0067]** GBS 80 refers to a putative cell wall surface anchor family protein. The nucleotide and amino acid sequences of GBS 80 sequenced from serotype V isolated strain 2603 V/R are set forth in Ref. 2 as SEQ ID 8779 and SEQ ID 8780. These sequences are also set forth below as SEQ ID NOS 1 and 2:

**SEQ ID NO. 1**

**[0068]** ATGAAATTATCGAAGAAGTTATTGTTTTCGGCTGCTGTTTTAACAATGGTG
GCGGGGTCAACTGTTGAACCAGTAGCTCAGTTTGCGACTGGAATGAGTATTGTAA
GAGCTGCAGAAGTGTCACAAGAACGCCCAGCGAAAACAACAGTAAATATCTATA
AATTACAAGCTGATAGTTATAAATCGGAAATTACTTCTAATGGTGGTATCGAGAA
TAAAGACGGCGAAGTAATATCTAACTATGCTAAACTTGGTGACAATGTAAAAGG
TTTGCAAGGTGTACAGTTTAAACGTTATAAAGTCAAGACGGATATTTCTGTTGAT
GAATTGAAAAAATTGACAACAGTTGAAGCAGCAGATGCAAAAGTTGGAACGATT
CTTGAAGAAGGTGTCAGTCTACCTCAAAAAACTAATGCTCAAGGTTTGGTCGTCG
ATGCTCTGGATTCAAAAAGTAATGTGAGATACTTGTATGTAGAAGATTTAAAGAA
TTCACCTTCAAACATTACCAAAGCTTATGCTGTACCGTTTGTGTTGGAATTACCAG
TTGCTAACTCTACAGGTACAGGTTTCCTTTCTGAAATTAATATTTACCCTAAAAAC
GTTGTAACTGATGAACCAAAAACAGATAAAGATGTTAAAAAATTAGGTCAGGAC
GATGCAGGTTATACGATTGGTGAAGAATTCAAATGGTTCTTGAAATCTACAATCC
CTGCCAATTTAGGTGACTATGAAAATTTGAAATTACTGATAAATTTGCAGATGGC

TTGACTTATAAATCTGTTGGAAAATCAAGATTGGTTCGAAAACACTGAATAGAGA
TGAGCACTACACTATTGATGAACCAACAGTTGATAACCAAAATACATTAAAAATT
ACGTTTAAACCAGAGAAATTTAAAGAAATTGCTGAGCTACTTAAAGGAATGACC
CTTGTTAAAAATCAAGATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGG
CATTTTTGGAAATTCCAGTTGCATCAACTATTAATGAAAAAGCAGTTTTAGGAAA
AGCAATTGAAAATACTTTTGAACTTCAATATGACCATACTCCTGATAAAGCTGAC
AATCCAAAACCATCTAATCCTCCAAGAAAACCAGAAGTTCATACTGGTGGGAAA
CGATTTGTAAAGAAAGACTCAACAGAAACACAAACACTAGGTGGTGCTGAGTTT
GATTTGTTGGCTTCTGATGGGACAGCAGTAAAATGGACAGATGCTCTTATTAAAG
CGAATACTAATAAAAACTATATTGCTGGAGAAGCTGTTACTGGGCAACCAATCA
AATTGAAATCACATACAGACGGTACGTTTGAGATTAAAGGTTTGGCTTATGCAGT
TGATGCGAATGCAGAGGGTACAGCAGTAACTTACAAATTAAAAGAAACAAAAGC
ACCAGAAGGTTATGTAATCCCTGATAAAGAAATCGAGTTTACAGTATCACAAAC
ATCTTATAATACAAAACCAACTGACATCACGGTTGATAGTGCTGATGCAACACCT
GATACAATTAAAAACAACAAACGTCCTTCAATCCCTAATACTGGTGGTATTGGTA
CGGCTATCTTTGTCGCTATCGGTGCTGCGGTGATGGCTTTTGCTGTTAAGGGGAT
GAAGCGTCGTACAAAAGATAAC

**SEQ ID NO: 2**

MKLSKKLLFS AAVLTMVAGS TVEPVAQFAT GMSIVRAAEV SQERPAKTTV
NIYKLQADSY KSEITSNGGI ENKDGEVISN YAKLGDNVKG LQGVQFKRYK 100
VKTDISVDEL KKLTTVEAAD AKVGTILEEG VSLPQKTNAQ GLVVDALDSK
SNVRYLYVED LKNSPSNITK AYAVPFVLEL PVANSTGTGF LSEINIYPKN 200
VVTDEPKTDK DVKKLGQDDA GYTIGEEFKW FLKSTIPANL GDYEKFEITD
KFADGLTYKS VGKIKIGSKT LNRDEHYTID EPTVDNQNTL KITFKPEKFK 300
EIAELLKGMT LVKNQDALDK ATANTDDAAF LEIPVASTIN EKAVLGKAIE
NTFELQYDHT PDKADNPKPS NPPRKPEVHT GGKRFVKKDS TETQTLGGAE 400
FDLLASDGTA VKWTDALIKA NTNKNYIAGE AVTGQPIKLK SHTDGTFEIK
GLAYAVDANA EGTAVTYKLK ETKAPEGYVI PDKEIEFTVS QTSYNTKPTD 500
ITVDSADATP DTIKNNKRPS *IPNTG*GIGTA IFVAIGAAVM AFAVKGMKRR
TKDN

**[0069]** GBS 80 contains an N-terminal leader or signal sequence region which is indicated by the underlined sequence at the beginning of SEQ ID NO: 2 above. GBS 80 also contains a

C-terminal transmembrane region which is indicated by the underlined sequence near the end of SEQ ID NO: 2 above. In preferred embodiments, the immunogenic polypeptides will have one or more amino acids from the transmembrane region and/or a cytoplasmic region removed to improve solubility of the antigen. GBS 80 contains an amino acid motif indicative of a cell wall anchor: **SEQ ID NO: 3** IPNTG (shown in italics in SEQ ID NO: 2 above). In some recombinant host cell systems, it may be preferable to remove this motif to facilitate secretion of a recombinant GBS 80 protein from the host cell. Accordingly, in preferred embodiments of the immunogenic polypeptides of GBS 80 for use in the invention, the transmembrane and/or cytoplasmic regions and the cell wall anchor motif are not included in the immunogenic polypeptides. Alternatively, in some recombinant host cell systems, it may be preferable to use the cell wall anchor motif to anchor the recombinantly expressed protein to the cell wall. The extracellular domain of the expressed protein may be cleaved during purification or the recombinant protein may be left attached to either inactivated host cells or cell membranes in the final composition.

**[0070]** A recombinantly produced GBS 80 fragment was used in the examples set out below that has the N-terminal leader sequence removed and replaced with an N-terminal methionine and the C-terminal cell wall anchor and transmembrane regions removed. The sequence of the recombinantly produced GBS 80 fragment is set out below:

```
MAEVSQERPA KTTVNIYKLQ ADSYKSEITS NGGIENKDGE VISNYAKLGD
NVKGLQGVQF KRYKVKTDIS VDELKKLTTV EAADAKVGTI LEEGVSLPQK 100
TNAQGLVVDA LDSKSNVRYL YVEDLKNSPS NITKAYAVPF VLELPVANST
GTGFLSEINI YPKNVVTDEP KTDKDVKKLG QDDAGYTIGE EFKWFLKSTI 200
PANLGDYEKF EITDKFADGL TYKSVGKIKI GSKTLNRDEH YTIDEPTVDN
QNTLKITFKP EKFKEIAELL KGMTLVKNQD ALDKATANTD DAAFLEIPVA 300
STINEKAVLG KAIENTFELQ YDHTPDKADN PKPSNPPRKP EVHTGGKRFV
KKDSTETQTL GGAEFDLLAS DGTAVKWTDA LIKANTNKNY IAGEAVTGQP 400
IKLKSHTDGT FEIKGLAYAV DANAEGTAVT YKLKETKAPE GYVIPDKEIE
FTVSQTSYNT KPTDITVDSA DATPDTIKNN KRPS (SEQ ID NO:4)
```

**[0071]** As described above, the invention includes fragments of a GBS 80 immunogenic polypeptide. The GBS 80 immunogenic polypeptides include the immunogenic epitopes of the cited GBS antigens may be used in the compositions of the invention.

**[0072]** Applicants have identified a particularly immunogenic fragment of the GBS 80 protein. This immunogenic fragment is located towards the N-terminus of the protein and is underlined in the GBS SEQ ID NO: 2 sequence below. The underlined fragment is set forth below as SEQ ID NO: 5.

**SEQ ID NO: 2**

MKLSKKLLFSAAVLTMVAGSTVEPVAQFATGMSIVRAAEVSQERPAKTTVNIYKLQ
ADSYKSEITSNGGIENKDGEVISNYAKLGDNVKGLQGVQFKRYKVKTDISVDELKKL
TTVEAADAKVGTILEEGVSLPQKTNAQGLVVDALDSKSNVRYLYVEDEKNSPSNITY
AVPFVLELPVANSTGTGFSEINIYPKNWTDEPKTDKDVKKLGQDDAGYTIGEEFKFK
STIPANLGDYEKFEITDKFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKI
TFKPEKFKEIAELLKGMTEVKNQDALDKATANTDDAAFLEIPVASTINEKAVLGKAIE
NTFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAEFDLLASD
GTAVKTDALIKANTNKNYIAGEAVTGQPIKKSHTDGTFEIKGLAYAVDANAEGTAV
TYKKETKAPEGYVIPDKEIEFTVSQTSYNTKPTDITVDSADATPDTIKNNKRPS*IPNTG*
GIGTAIFVAIGAAVMAFAVKGMKRRTKDN

**SEQ ID NO: 5**

AEVSQERPAKTTVNIYKLQADSYKSEITSNGGIENKDGEVISNYAKLGDNVK**GLQGV**
**QFKRYKVKTDISVDELKKLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVDAL**
**DSKSNVR**YLYVEDLKNSPSNITKAYAVPFVLELPVANSTGTGFLSEINIYPKNWTDEP
KTDKDVKKLGQDDAGYTIGEEFKWFLKSTIPANLGDYEKFEITDKFADGLTYKSVGK
IKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFKEIAELLKG

**[0073]** Two of the immunogenic polypeptides identified in Example 2 are shown in SEQ ID NO:5 above. SEQ ID NO: 7 is underlined and SEQ ID NO: 8 is highlighted in bold.

**SEQ ID NO: 7**

DGEVISNYAKLGDNVKGLQGVQFKRYKVKTDISVDELKKLTTVEAADAKVGTILEE
GVSLPQKTNAQGLVVDALDSKSNVR

**SEQ ID NO: 8**

GLQGVQFKRYKVKTDISVDELKKLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVD ALDSKSNVR

**[0074]** The immunogenicity of the protein encoded by SEQ ID NO: 5 was compared against PBS, GBS whole cell, GBS 80 (full length) and another fragment of GBS 80, located closer to the C terminus of the peptide (SEQ ID NO: 6, below).

**SEQ ID NO: 6**

MTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVLGKAIENTFELQYDGTPDKA DNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAEFDLLASDGTAVKWTDALIKA NTNKNYIAGEAVTGQPIKLKSHTDGTFEIKGLAYAVDANAEGTAVTYKLKETIAPEG YVIPDKEIEFTVSQTSYNTKPTDITVDSADATPDTIKNNKRPS

**[0075]** Two of the immunogenic polypeptides identified in Example 2 are shown in SEQ ID NO: 6 above. SEQ ID NO: 9 is underlined and SEQ ID NO: 10 is highlighted in bold.

**SEQ ID NO: 9**

YDGTPDKADNPKPSNPPRKPEVHTGGKRFV

**SEQ ID NO: 10**

NPKPSNPPR

**[0076]** The peptide array epitope mapping described in Example 3 identified two additional immunogenic polypeptides – DALDSKSNVRYLY (SEQ ID NO:11) and SNVRYLYVEDLKN (SEQ ID NO:12).

GBS 80 Immunogenic polypeptides

**[0077]** As discussed above, one embodiment of the invention provides an immunogenic composition comprising an immunogenic polypeptide from GBS 80 or a fragment thereof, wherein said immunogenic polypeptide is a fragment of GBS 80 that includes one of the regions identified in this application and may include additional portions of GBS 80. Of particular interest are the immunogenic polypeptides of SEQ ID NOs: 7-12. The length of the fragment may vary depending on the amino acid sequence of the specific immunogenic

polypeptide, but the fragment is preferably at least 7 consecutive amino acids, (*e.g.* 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 or more).

**[0078]** The immunogenic polypeptides may include polypeptide sequences having sequence identity to the identified immunogenic polypeptides. The degree of sequence identity may vary depending on the amino acid sequence in question, but is preferably greater than 50% (*e.g.* 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5% or more). Polypeptides having sequence identity include homologs, orthologs, allelic variants and functional mutants of the identified GBS 80 immunogenic polypeptides. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence. Identity between proteins is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), using an affinity gap search with parameters *gap open penalty=12* and *gap extension penalty=1*. The immunogenic polypeptides may include polypeptide sequences encoded by nucleic acid sequences that hybridize under high stringency wash conditions (see below for representative conditions) to nucleic acids encoding an identified immunogenic polypeptide (especially SEQ ID NO: 7-12 or an antigenic fragment thereof).

**[0079]** With regard to the immunogenic polypeptide of SEQ ID NO:7, polypeptides of particular interest include polypeptides having limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 259 or fewer amino acids to SEQ ID NO:2 including SEQ ID NO:7. In preferred embodiments, the contiguous region will be "over 250," "over 240," "over 230," "over 220," "over 210," "over 200," "over 180," "over 160," "over 140," "over 120," "over 100," "over 80," "over 60," "over 50," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14," "over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous identity is desired and the phrase "or fewer" may be replaced with "to 200," "to 180," "to 160," "to 140," "to 120," "to 100," "to 80," "to 60," "to 50," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0080]** With regard to the immunogenic polypeptide of SEQ ID NO:8, polypeptides of particular interest include polypeptides having limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 259 or fewer amino acids to SEQ ID NO:2 including SEQ ID NO:8. In preferred embodiments, the contiguous region will be "over 250," "over 240," "over 230," "over 220," "over 210," "over 200," "over 180," "over 160," "over 140," "over 120," "over 100," "over 80," "over 60," "over 50," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14," "over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous identity is desired and the phrase "or fewer" may be replaced with "to 200," "to 180," "to 160," "to 140," "to 120," "to 100," "to 80," "to 60," "to 50," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0081]** With regard to the immunogenic polypeptide of SEQ ID NO:9, polypeptides of particular interest include polypeptides having limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 148 or fewer amino acids to SEQ ID NO:2 including SEQ ID NO:9. In preferred embodiments, the contiguous region will be "over 147," "over 146," "over 145," "over 140," "over 130," "over 120," "over 100," "over 80," "over 74," "over 72," "over 70," "over 65," "over 60," "over 55," "over 50," "over 45," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14," "over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous identity is desired and the phrase "or fewer" may be replaced with "to 147," "to 146," "to 145," "to 140," "to 130," "to 120," "to 100," "to 80," "to 74," "to 72," "to 70," "to 65," "to 60," "to 55," "to 50," "to 45," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0082]** With regard to the immunogenic polypeptide of SEQ ID NO:10, polypeptides of particular interest include polypeptides having limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) over 76 or fewer amino acids to SEQ ID NO:2 including SEQ ID NO:10. In preferred embodiments, the contiguous region will be "over 74," "over 72," "over 70," "over 65," "over 60," "over 55," "over 50," "over 45," "over 40," "over 35," "over 30," "over 27," "over 23," "over 20," "over 18," "over 16," "over 14," "over 13," "over 11," "over 10," "over 9," "over 8," or "over 7." In certain embodiments, a lower limit on the region of contiguous identity is desired and the phrase "or fewer" may be replaced with "to 70," "to 65," "to 60," "to 55," "to 50," "to 45," "to 40," "to 35," "to 30," "to 27," "to 23," "to 20," "to 18," "to 16," "to 14," "to 13," "to 11," "to 10," "to 9," "to 8," or "to 7." One of skill in the art will appreciate that any pair-wise combination of limits may be selected as desired and that the same upper and lower limit may be selected to specify the desired length of the region of contiguous alignment.

**[0083]** With regard to the immunogenic polypeptide of SEQ ID NO:9, additional polypeptides of particular interest include polypeptides having limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) percent to SEQ ID NO:2 including SEQ ID NO:9 extending no more than 47 amino acids upstream. In preferred embodiments, the region of contiguous alignment will extend no more than 46, 44, 42, 40, 35, 30, 25, 20, 15, 10, 8, 7, 5, 4, 3, 2, or 1 amino acid(s) upstream. In some embodiments, the region of contiguous alignment will begin with the N-terminal end of SEQ ID NO:9.

**[0084]** With regard to the immunogenic polypeptide of SEQ ID NO:10, additional polypeptides of particular interest include polypeptides having limited, contiguous sequence identity of at least 50 percent (or with increasing preference at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5%) percent to SEQ ID NO:2 including SEQ ID NO:10 extending no more than 56 amino acids upstream. In preferred embodiments, the region of contiguous alignment will extend no more than 54, 52, 50, 48, 46, 44, 42, 40, 35, 30, 25, 20, 15, 10, 8, 7,

5, 4, 3, 2, or 1 amino acid(s) upstream. In some embodiments, the region of contiguous alignment will begin with the N-terminal end of SEQ ID NO:10.

Expression systems

**[0085]** The GBS 80 immunogenic polypeptide nucleotide sequences can be expressed in a variety of different expression systems; for example those used with mammalian cells, baculoviri, plants, bacteria, and yeast.

*i. Mammalian Systems*

**[0086]** Mammalian expression systems are known in the art. A mammalian promoter is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiating region, which is usually placed proximal to the 5' end of the coding sequence, and a TATA box, usually located 25- 30 base pairs (bp) upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase 11 to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element, usually located within 100 to 200 bp upstream of the TATA box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation (Sambrook et. (1989) *Expression of Cloned Genes in Mammalian Cells.* In Molecular Cloning: A Laboratory Manual, 2nd ed.).

**[0087]** Mammalian viral genes are often highly expressed and have a broad host range; therefore sequences encoding mammalian viral genes provide particularly useful promoter sequences. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter (Ad MLP), and herpes simplex virus promoter. In addition, sequences derived from non-viral genes, such as the murine metallothionein gene, also provide useful promoter sequences, Expression may be either constitutive or regulated (inducible), depending on the promoter can be induced with glucocorticoid in hormone-responsive cells.

**[0088]** The presence of an enhancer element (enhancer), combined with the promoter elements described above, will usually increase expression levels. An enhancer is a regulatory DNA sequence that can stimulate transcription up to 1000-fold when linked to homologous or heterologous promoters, with synthesis beginning at the normal RNA start

site. Enhancers are also active when they are placed upstream or downstream from the transcription initiation site, in either normal or flipped orientation, or at a distance of more than 1000 nucleotides from the promoter (Maniatis et al. (1987) Science 236:1237; Alberts et al. (1989) Molecular Biology of the Cell, 2nd ed.). Enhancer elements derived from viruses may be particularly useful, because they usually have a broader host range. Examples include the SV40 early gene enhancer (Dijkema et al (1985) EMBO J. 4:7611) and the enhancer/promoters derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus (Gorman et al. (1982b) Proc. Natl. Acad. Sci. 79:6777) and from human cytomegalovirus (Boshart et al. (1985) Cell 41:5211). Additionally, some enhancers are regulatable and become active only in the presence of an inducer, such as a hormone or metal ion (Sassone-Corsi and Borelli (1986) Trends Genet. 2:215; Maniatis et al. (1987) Science 236:1237).

**[0089]** A DNA molecule may be expressed intracellularly in mammalian cells. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, the N-terminus may be cleaved from the protein by in vitro incubation with cyanogen bromide.

**[0090]** Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in mammalian cells, Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either in vivo or in vitro. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The adenovirus tripartite leader is an example of a leader sequence that provides for secretion of a foreign protein in mammalian cells.

**[0091]** Usually, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3' terminus of the mature mRNA is formed by site-specific post-transcriptional cleavage and polyadenylation (Birnstiel et al. (1985) Cell 41:349; Proudfoot and Whitelaw (1988) *Termination and 3' end processing of eukaryotic RNA*. In Transcription and splicing (ed. B.D. Hames and D.M, Glover); Proudfoot (1989) Trends Biochem. Sci. 14:1051). These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA.

Examples of transcription terminater/polyadenylation signals include those derived from SV40 (Sambrook et al (1989) *Expression of cloned genes in cultured mammalian cells.* In Molecular Cloning: A Laboratory Manual).

**[0092]** Usually, the above described components, comprising a promoter, polyadenylation signal, and transcription termination sequence are put together into expression constructs. Enhancers, introns with functional splice donor and acceptor sites, and leader sequences may also be included in an expression construct, if desired. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as mammalian cells or bacteria. Mammalian replication systems include those derived from animal viruses, which require trans-acting factors to replicate. For example, plasmids containing the replication systems of papovaviri, such as SV40 (Gluzman (1981) Cell 23:1751) or polyomavirus, replicate to extremely high copy number in the presence of the appropriate viral T antigen. Additional examples of mammalian replicons include those derived from bovine papillomavirus and Epstein- Barf virus. Additionally, the replicon may have two replication systems, thus allowing it to be maintained, for example, in mammalian cells for expression and in a prokaryotic host for cloning and amplification. Examples of such mammalian-bacteria shuttle vectors include pMT2 (Kaufman et al. (1989) Mol. Cell. Biol. 9:946) and pHEBO (Shimizu et al. (1986) Mol. Cell. Biol. 6:10741). The transformation procedure used depends upon the host to be transformed. Methods for introduction of heterologous polynucleotides into mammalian cells are known in the art and include dextran-mediated transfection, calcium phosphate precipitation, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in 11posomes, and direct microinjection of the DNA into nuclei.

**[0093]** Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e. g., Hep G2), and a number of other cell lines.

*ii. Baculovirus Systems*

**[0094]** The polynucleotide encoding the protein can also be inserted into a suitable insect expression vector, and is operably linked to the control elements within that vector. Vector

construction employs techniques which are known in the art. Generally, the components of the expression system include a transfer vector, usually a bacterial plasmid, which contains both a fragment of the baculovirus genome, and a convenient restriction site for insertion of the heterologous gene or genes to be expressed; a wild type baculovirus with a sequence homologous to the baculovirus- specific fragment in the transfer vector (this allows for the homologous recombination of the heterologous gene in to the baculovirus genome); and appropriate insect host cells and growth media, After inserting the DNA sequence encoding the protein into the transfer vector, the vector and the wild type viral genome are transfected into an insect host cell where the vector and viral genome are allowed to recombine. The packaged recombinant virus is expressed and recombinant plaques are identified and purified. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, inter alia, Invitrogen, San Diego CA ("MaxBac" kit). These techniques are generally known to those skilled in the art and fully described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987) (hereinafter "Summers and Smith").

**[0095]** Prior to inserting the DNA sequence encoding the protein into the baculovirus genome, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are usually assembled into an intermediate transplacement construct (transfer vector). This construct may contain a single gene and operably linked regulatory elements; multiple genes, each with its owned set of operably linked regulatory elements; or multiple genes, regulated by the same set of regulatory elements. Intermediate transplacement constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as a bacterium, The replicon will have a replication system, thus allowing it to be maintained in a suitable host for cloning and amplification.

**[0096]** Currently, the most commonly used transfer vector for introducing foreign genes into AcNPV is pAc373. Many other vectors, known to those of skill in the art, have also been designed. These include, for example, pVL985 (which alters the polyhedrin start codon from ATG to ATT, and which introduces a BamHI cloning site 32 basepairs downstream from the ATT (Luckow and Summers, Virology (1989) 17:31).

**[0097]** The plasmid usually also contains the polyhedrin polyadenylation signal (Miller et al. (1988) Ann. Rev. Microbiol., 42:177) and a prokaryotic ampicillin-resistance (amp) gene and origin of replication for selection and propagation in E. coli.

**[0098]** Baculovirus transfer vectors usually contain a baculovirus promoter. A baculovirus promoter is any DNA sequence capable of binding a baculovirus RNA polymerase and initiating the downstream (5' to 3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A baculovirus transfer vector may also have a second domain called an enhancer, which, if present, is usually distal to the structural gene. Expression may be either regulated or constitutive.

**[0099]** Structural genes, abundantly transcribed at late times in a viral infection cycle, provide particularly useful promoter sequences. Examples include sequences derived from the gene encoding the viral polyhedron protein (Friesen et al., (1986) *The Regulation of Baculovirus Gene Expression,* in: The Molecular Biology of Baculoviruses (ed. Walter Doerfler); EPO Publ. Nos. 127 839 and 155 476) and the gene encoding the p10 protein (Vlak et al, (1988), J. Gen. Virol. 69:765).

**[0100]** DNA encoding suitable signal sequences can be derived from genes for secreted insect or baculovirus proteins, such as the baculovirus polyhedrin gene (Carbonell et al. (1988) Gene, 73:409). Alternatively, since the signals for mammalian cell post-translational modifications (such as signal peptide cleavage, proteolytic cleavage, and phosphorylation) appear to be recognized by insect cells, and the signals required for secretion and nuclear accumulation also appear to be conserved between the invertebrate cells and vertebrate cells, leaders of non-insect origin, such as those derived from genes encoding human α-interferon (Maeda et al., (1985), Nature 315:592); human gastrin-releasing peptide (Lebacq-Verheyden et al., (1988), Molec. Cell. Biol. 8:3129); human IL-2 (Smith et al., (1985) Proc. Nat'l Acad. Sci. USA, 82:8404); mouse IL-3 (Miyajima et al., (1987) Gene 58:273); and human glucocerebrosidase (Martin et al. (1988) DNA, 7:99), can also be used to provide for secretion in insects.

**[0101]** A recombinant polypeptide or polyprotein may be expressed intracellularly or, if it is expressed with the proper regulatory sequences, it can be secreted. Good intracellular

expression of non-fused foreign proteins usually requires heterologous genes that ideally have a short leader sequence containing suitable translation initiation signals preceding an ATG start signal. If desired, methionine at the N-terminus may be cleaved from the mature protein by in vitro incubation with cyanogen bromide.

**[0102]** Alternatively, recombinant polyproteins or proteins which are not naturally secreted can be secreted from the insect cell by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in insects. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the translocation of the protein into the endoplasmic reticulum.

**[0103]** After insertion of the DNA sequence and/or the gene encoding the expression product precursor of the protein, an insect cell host is co- transformed with the heterologous DNA of the transfer vector and the genomic DNA of wild type baculovirus -- usually by co-transfection. The promoter and transcription termination sequence of the construct will usually comprise a 2-5 kb section of the baculovirus genome. Methods for introducing heterologous DNA into the desired site in the baculovirus virus are known in the art, (See Summers and Smith supra; Ju et al. (1987); Smith et al., Mol. Cell. Biol. (1983) 3:2156; and Luckow and Summers (1989)). For example, the insertion can be into a gene such as the polyhedrin gene, by homologous double crossover recombination; insertion can also be into a restriction enzyme site engineered into the desired baculovirus gene (Miller et al., (1989), Bioessays 4:91). The DNA sequence, when cloned in place of the polyhedrin gene in the expression vector, is flanked both 5' and 3' by polyhedrin-specific sequences and is positioned downstream of the polyhedrin promoter.

**[0104]** The newly formed baculovirus expression vector is subsequently packaged into an infectious recombinant baculovirus. Homologous recombination occurs at low frequency (between about 1% and about 5%); thus, the majority of the virus produced after cotransfection is still wild-type virus. Therefore, a method is necessary to identify recombinant viruses. An advantage of the expression system is a visual screen allowing recombinant viruses to be distinguished. The polyhedrin protein, which is produced by the native virus, is produced at very high levels in the nuclei of infected cells at late times after viral infection. Accumulated polyhedrin protein forms occlusion bodies that also contain embedded particles. These occlusion bodies, up to 15 $\mu$m in size, are highly refractile, giving

them a bright shiny appearance that is readily visualized under the light microscope. Cells infected with recombinant viruses lack occlusion bodies. To distinguish recombinant virus from wildtype virus, the transfection supernatant is plaqued onto a monolayer of insect cells by techniques known to those skilled in the 'art. Namely, the plaques are screened under the light microscope for the presence (indicative of wild-type virus) or absence (indicative of recombinant virus) of occlusion bodies (*Current Protocols in Microbiology*, Vol. 2 (Ausubel et al. eds) at 16.8 (Supp. 10, 1990); Summers and Smith, supra; Miller et al. (1989)).

**[0105]** Recombinant baculovirus expression vectors have been developed for infection into several insect cells. For example, recombinant baculoviruses have been developed for, inter alia: *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni* (WO 89/046699; Carbonell et al., (1985) J. Virol. 56:153; Wright (1986) Nature 321:718; Smith et al., (1983) Mol. Cell. Biol. 3:2156; and see generally, Fraser, et al. (1989) In Vitro Cell. Dev. Biol. 25:225).

**[0106]** Cells and cell culture media are commercially available for both direct and fusion expression of heterologous polypeptides in a baculovirus/expression system; cell culture technology is generally known to those skilled in the art. See, e.g., Summers and Smith supra.

**[0107]** The modified insect cells may then be grown in an appropriate nutrient medium, which allows for stable maintenance of the plasmid(s) present in the modified insect host. Where the expression product gene is under inducible control, the host may be grown to high density, and expression induced. Alternatively, where expression is constitutive, the product will be continuously expressed into the medium and the nutrient medium must be continuously circulated, while removing the product of interest and augmenting depleted nutrients. The product may be purified by such techniques as chromatography, e.g., HPLC, affinity chromatography, ion exchange chromatography, etc.; electrophoresis; density gradient centrifugation; solvent extraction, or the like. As appropriate, the product may be further purified, as required, so as to remove substantially any insect proteins which are also secreted in the medium or result from lysis of insect cells, so as to provide a product which is at least substantially free of host debris, e.g., proteins, lipids and polysaccharides.

**[0108]** In order to obtain protein expression, recombinant host cells derived from the transformants are incubated under conditions which allow expression of the recombinant

protein encoding sequence. These conditions will vary, dependent upon the host cell selected. However, the conditions are readily ascertainable to those of ordinary skill in the art, based upon what is known in the art.

*iii. Plant Systems*

**[0109]** There are many plant cell culture and whole plant genetic expression systems known in the art. Exemplary plant cellular genetic expression systems include those described in patents, such as: US 5,693,506; US 5, 659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture have been described by Zenk, Phytochemistry 30:3861- 3863 (1991). Descriptions of plant protein signal peptides may be found in addition to the references described above in Vaulcombe et al., Mol. Gen. Genet. 209:33-40 (1987); Chandler et al., Plant Molecular Biology 3:407-418 (1984); Rogers, J. Biol. Chem. 260:3731-3738 (1985); Rothstein et al., Gene 55:353-356 (1987); Whittier et al., Nucleic Acids Research 15:2515-2535 (1987); Wirsel et al., Molecular Microbiology 3:3-14 (1989); and Yu et al., Gene 122:247-253 (1992). A description of the regulation of plant gene expression by the phytohormone, gibberellic acid and secreted enzymes induced by gibberellic acid can be found in R.L. Jones and J. MacMillin, Gibberellins: in: Advanced Plant Physiology, Malcolm B. Wilkins, ed., 1984 Pitman Publishing Limited, London, pp. 21-52. References that describe other metabolically-regulated genes: Sheen, Plant Cell, 2:1027-1038 (1990); Maas et al., EMBO J. 9:3447-3452 (1990); Benkel and Hickey, Proc. Natl. Acad. Sci. 84:1337-1339 (1987)

**[0110]** Typically, using techniques known in the art, a desired polynucleotide sequence is inserted into an expression cassette comprising genetic regulatory elements designed for operation in plants. The expression cassette is inserted into a desired expression vector with companion sequences upstream and downstream from the expression cassette suitable for expression in a plant host. The companion sequences will be of plasmid or viral origin and provide necessary characteristics to the vector to permit the vectors to move DNA from an original cloning host, such as bacteria, to the desired plant host. The basic bacterial/plant vector construct will preferably provide a broad host range prokaryote replication origin; a prokaryote selectable marker; and, for Agrobacterium transformations, T DNA sequences for Agrobacterium-mediated transfer to plant chromosomes. Where the heterologous gene is not readily amenable to detection, the construct will preferably also have a selectable marker gene suitable for determining if a plant cell has been transformed. A general review of

suitable markers, for example for the members of the grass family, is found in Wilmink and Dons, 1993, Plant Mol. Biol. Reptr, 11 (2):165-185.

**[0111]** Sequences suitable for permitting integration of the heterologous sequence into the plant genome are also recommended. These might include transposon sequences and the like for homologous recombination as well as Ti sequences which permit random insertion of a heterologous expression cassette into a plant genome. Suitable prokaryote selectable markers include resistance toward antibiotics such as ampicillin or tetracycline. Other DNA sequences encoding additional functions may also be present in the vector, as is known in the art.

**[0112]** The nucleic acid molecules of the subject invention may be included into an expression cassette for expression of the protein(s) of interest. Usually, there will be only one expression cassette, although two or more are feasible. The recombinant expression cassette will contain in addition to the heterologous protein encoding sequence the following elements, a promoter region, plant 5' untranslated sequences, initiation codon depending upon whether or not the structural gene comes equipped with one, and a transcription and translation termination sequence. Unique restriction enzyme sites at the 5' and 3' ends of the cassette allow for easy insertion into a pre-existing vector.

**[0113]** A heterologous coding sequence may be for any protein relating to the present invention. The sequence encoding the protein of interest will encode a signal peptide which allows processing and translocation of the protein, as appropriate, and will usually lack any sequence which might result in the binding of the desired protein of the invention to a membrane. Since, for the most part, the transcriptional initiation region will be for a gene which is expressed and translocated during germination, by employing the signal peptide which provides for translocation, one may also provide for translocation of the protein of interest. In this way, the protein(s) of interest will be translocated from the cells in which they are expressed and may be efficiently harvested. Typically secretion in seeds are across the aleurone or scutellar epithelium layer into the endosperm of the seed. While it is not required that the protein be secreted from the cells in which the protein is produced, this facilitates the isolation and purification of the recombinant protein.

**[0114]** Since the ultimate expression of the desired gene product will be in a eukaryotic cell it is desirable to determine whether any portion of the cloned gene contains sequences which will be processed out as introns by the host's splicosome machinery. If so, site-directed

mutagenesis of the "intron" region may be conducted to prevent losing a portion of the genetic message as a false intron code (Reed and Maniatis, Cell 41:95-105, 1985).

**[0115]** The vector can be microinjected directly into plant cells by use of micropipettes to mechanically transfer the recombinant DNA (Crossway, Mol. Gen. Genet, 202:179-185, 1985). The genetic material may also be transferred into the plant cell by using polyethylene glycol (Krens, et al., Nature, 296, 72-74, 1982). Another method of introduction of nucleic acid segments is high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface (Klein, et al., Nature, 327, 70-73, 1987 and Knudsen and Muller, 1991, Planta, 185:330-336) teaching particle bombardment of barley endosperm to create transgenic barley. Yet another method of introduction would be fusion of protoplasts with other entities, either minicells, cells, lysosomes or other fusible lipid-surfaced bodies, Fraley, et al., Proc. Natl. Acad. Sci. USA, 79, 1859-1863, 1982.

**[0116]** The vector may also be introduced into the plant cells by electroporation (Fromm et al., Proc. Natl Acad. Sci. USA 82:5824, 1985). In this technique, plant protoplasts are electroporated in the presence of plasmids containing the gene construct. Electrical impulses of high field strength reversibly permeablize membranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and form plant callus.

**[0117]** All plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred gene. It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Some suitable plants include, for example, species from the genera *Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersion, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Cichorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Hererocallis, Nemesia, Pelargonium, Panicum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Lolium, Zea, Triticum, Sorghum,* and *Datura.*

**[0118]** Means for regeneration vary from species to species of plants, but generally a suspension of transformed protoplasts containing copies of the heterologous gene is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced from the protoplast suspension. These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Shoots and roots normally develop simultaneously. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is fully reproducible and repeatable.

**[0119]** In some plant cell culture systems, the desired protein of the invention may be excreted or alternatively, the protein may be extracted from the whole plant. Where the desired protein of the invention is secreted into the medium, it may be collected. Alternatively, the embryos and embryoless-half seeds or other plant tissue may be mechanically disrupted to release any secreted protein between cells and tissues. The mixture may be suspended in a buffer solution to retrieve soluble proteins. Conventional protein isolation and purification methods will be then used to purify the recombinant protein. Parameters of time, temperature pH, oxygen, and volumes will be adjusted through routine methods to optimize expression and recovery of heterologous protein.

*iv. Bacterial Systems*

**[0120]** Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g., a structural gene) into mRNA, A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the

RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps initiate transcription of the lac operon in *Escherichia coli* (Raibaud et al. (1984) Annu. Rev. Genet. 18:173). Regulated expression may therefore either be positive or negative, thereby either enhancing or reducing transcription.

**[0121]** Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (lac) (Chang et al. (1977) Nature 198:1056), and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (trp) (Goeddel et al. (1980) Nuc. Acids Res. 8:4057; Yelverton et al. (1981) Nucl. Acids Res. 9:731; US patent 4,738, 921; EP-A-0036776 and EP-A- 0121775); and the β-lactamase (bla) promoter system (Weissmann (1981) "The cloning of interferon and other mistakes." In Interferon 3 (ed. 1. Gresser)). The bacteriophage lambda PL (Shimatake et al. (1981) Nature 292:128) and T5 (US patent 4,689,406) promoter systems also provide useful promoter sequences.

**[0122]** In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter (US patent 4,551,4331). For example, the tac promoter is a hybrid trp-lac promoter comprised of both trp promoter and lac operon sequences that is regulated by the lac repressor (Amann et al. (1983) Gene 25:167; de Boer et al. (1983) Proc. Natl. Acad. Sci. 80:21). Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophage T7 RNA polymerase/promoter system is an example of a coupled promoter system (Studier et al. (1986) J. Mol. Biol. 189:113; Tabor et al. (1985) Proc Natl. Acad. Sci. 82:1074). In addition, a hybrid promoter can also be comprised of a bacteriophage promoter and an E. coli operator region (EPO-A-0 267 851).

**[0123]** In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In *E. coli*, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon

(ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon (Shine et al. (1975) Nature 254:34). The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' end of E. coli 16S rRNA (Steitz et al. (1979) "Genetic signals and nucleotide sequences in messenger RNA." In Biological Regulation and Development: Gene Expression (ed. R.F. Goldberger)). To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site (Sambrook et al. (1989) "Expression of cloned genes in Escherichia coli." In Molecular Cloning: A Laboratory Manual).

**[0124]** A DNA molecule may be expressed intracellularly. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by in vitro incubation with cyanogen bromide or by either in vivo on in vitro incubation with a bacterial methionine N-terminal peptidase (EPO-A-0 219 237).

**[0125]** Fusion proteins provide an alternative to direct expression. Usually, a DNA sequence encoding the N-terminal portion of an endogenous bacterial protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the bacteriophage lambda cell gene can be linked at the 5' terminus of a foreign gene and expressed in bacteria. The resulting fusion protein preferably retains a site for a processing enzyme (factor Xa) to cleave the bacteriophage protein from the foreign gene (Nagai et al. (1984) Nature 309:8101). Fusion proteins can also be made with sequences from the lacZ (Jia et al. (1987) Gene 60:197), trpE (Allen et al. (1987) J. Biotechnol. 5:93; Makoff et al. (1989) J. Gen. Microbiol. 135:11), and Chey (EP-A-0 324 647) genes. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g., ubiquitin specific processing-protease) to cleave the ubiquitin from the foreign protein. Through this method, native foreign protein can be isolated (Miller et al. (1989) Bio/Technology 7:698).

**[0126]** Alternatively, foreign proteins can also be secreted from the cell by creating chimeric DNA molecules that encode a fusion protein comprised of a signal peptide sequence fragment that provides for secretion of the foreign protein in bacteria (US patent 4,336,336).

The signal sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The protein is either secreted into the growth media (gram-positive bacteria) of into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). Preferably there are processing sites, which can be cleaved either in vivo or in vitro encoded between the signal peptide fragment and the foreign gene.

**[0127]** DNA encoding suitable signal sequences can be derived from genes for secreted bacterial proteins, such as the E. coli outer membrane protein gene (ompA) (Masui et al. (1983), in: Experimental Manipulation of Gene Expression; Ghrayeb et al. (1984) EMBO J. 3:2437) and the E. coli alkaline phosphatase signal sequence (phoA) (Oka et al. (1985) Proc. Natl. Acad. Sci. 82:7212). As an additional example, the signal sequence of the alpha-amylase gene from various Bacillus strains can be used to secrete heterologous proteins from B. subtilis (Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; EP-A-0 244 042).

**[0128]** Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Transcription termination sequences frequently include DNA sequences of about 50 nucleotides capable of forming stem loop structures that aid in terminating transcription. Examples include transcription termination sequences derived from genes with strong promoters, such as the trp gene in E, coli as well as other biosynthetic genes.

**[0129]** Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a prokaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy

number vector may be selected, depending upon the effect of the vector and the foreign protein on the host.

**[0130]** Alternatively, the expression constructs can be integrated into the bacterial genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to the bacterial chromosome that allows the vector to integrate. Integrations appear to result from recombinations between homologous DNA in the vector and the bacterial chromosome, For example, integrating vectors constructed with DNA from various Bacillus strains integrate into the Bacillus chromosome (EP-A- 0 127 328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

**[0131]** Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of bacterial strains that have been transformed. Selectable markers can be expressed in the bacterial host and may include genes which render bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin (neomycin), and tetracycline (Davies et al. (1978) Annu. Rev. Microbiol. 32:469). Selectable markers may also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways.

**[0132]** Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable market that is either maintained in a replicon or developed into an integrating vector, as described above.

**[0133]** Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For example, expression vectors have been developed for, inter alia, the following bacteria: Bacillus subtilis (Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; EP-A-0 036 259 and EP-A-0 063 953; WO 84/04541), Escherichia coli (Shimatake et al. (1981) Nature 292:128; Amann et al. (1985) Gene 40:183; Studier et al. (1986) J. Mol. Biol. 189:113; EP-A-0 036 776,EPA-0 136 829 and EP-A-0 136 907), Streptococcus cremoris (Powell et al. (1988) Appl. Environ. Microbiol. 54:655), Streptococcus lividans (Powell et al. (1988) Appl. Environ. Microbiol. 54:655), and Streptomyces lividans (US patent 4,745,056).

**[0134]** Methods of introducing exogenous DNA into bacterial hosts are well- known in the art, and usually include either the transformation of bacteria treated with CaC12 or other

agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by electroporation. Transformation procedures usually vary with the bacterial species to be transformed. See e.g., (Masson et al. (1989) FEMS Microbiol. Lett. 60:273; Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; EP-A-0 036 259 and EP-A-0 063 953; WO 84/04541, Bacillus) (Miller et al. (1988) Proc. Natl. Acad. Sci. 85:856; Wang et al. (1990) J. Bacteriol. 172:949, Campylobacter), (Cohen et al. (1973) Proc. Natl. Acad. Sci. 69:2110; Dower et al. (1988) Nucleic Acids Res. 16:6127; Kushner (1978) "An improved method for transformation of Escherichia coli with ColEl-derived plasmids. In Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering (eds. H.W. Boyer and S. Nicosia); Mandel et al. (1970) J. Mol. Biol. 53:159; Taketo (1988) Biochim. Biophys. Acta 949:318; Escherichia), (Chassy et al. (1987) FEMS Microbiol. Lett. 44:173, Lactobacillus), (Fiedler et al. (1988) Anal. Biochem 170:38, Pseudomonas), (Augustin et al. (1990) FEMS Microbiol. Lett. 66:203, Staphylococcus), (Barany et al. (1980) J. Bacteriol. 144:698; Harlander (1987) "Transformation of Streptococcus lactis by electroporation, in: Streptococcal Genetics (ed. J. Ferretti and R. Curtiss 111); Perry et al. (1981) Infect. Immun. 32:1295; Powell et al. (1988) Appl. Environ. Microbiol. 54:655; Somkuti et al. (1987) Proc. 4th Evr. Cong. Biotechnology 1:412, Streptococcus).

*v. Yeast Expression*

**[0135]** Yeast expression systems are also known to one of ordinary skill in the art. A yeast promoter is any DNA sequence capable of binding yeast RNA polymerase and initiating the downstream 3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site (the "TATA Box") and a transcription initiation site. A yeast promoter may also have a second domain called an upstream activator sequence (UAS), which, if present, is usually distal to the structural gene. The UAS permits regulated (inducible) expression. Constitutive expression occurs in the absence of a UAS, but may be enhanced with one or more UAS. Regulated expression may be either positive or negative, thereby either enhancing or reducing transcription.

**[0136]** Yeast is a fermenting organism with an active metabolic pathway, therefore sequences encoding enzymes in the metabolic pathway provide particularly useful promoter sequences. Examples include alcohol dehydrogenase (ADH) (EP-A-0 284 044), enolase,

glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (EPO-A-0 329 203). The yeast PH05 gene, encoding acid phosphatase, also provides useful promoter sequences (Myanohara et al. (1983) Proc. Natl. Acad. Sci. USA 80:1).

**[0137]** In addition, synthetic promoters which do not occur in nature also function as yeast promoters. For example, UAS sequences of one yeast promoter may be joined with the transcription activation region of another yeast promoter, creating a synthetic hybrid promoter. Examples of such hybrid promoters include the ADH regulatory sequence linked to the GAP transcription activation region (US Patent Nos. 4,876,197 and 4,880, 734). Other examples of hybrid promoters include promoters which consist of the regulatory sequences of either the AD112, GAL4, GALIO, OR PH05 genes, combined with the transcriptional activation region of a glycolytic enzyme gene such as GAP or PyK (EP-A-0 164 556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, inter alia, (Cohen et al. (1980) Proc. Natl. Acad. Sci. USA 77:1078; Henikoff et al. (1981) Nature 283:835; Hollenberg et al. (1981) Curr. Topics Microbiol. Immunol. 96:119; Hollenberg et al. (1979) "The Expression of Bacterial Antibiotic Resistance Genes in the Yeast Saccharomyces cerevisiae," in: Plasmids of Medical, Environmental and Commercial Importance (eds. K.N. Timmis and A. Puhler); Mercerau-Puigalon et al. (1980) Gene 11:163; Panthier et al. (1980) Curr. Genet. 2:109).

**[0138]** A DNA molecule may be expressed intracellularly in yeast. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by in vitro incubation with cyanogen bromide.

**[0139]** Fusion proteins provide an alternative for yeast expression systems, as well as in mammalian, baculovirus, and bacterial expression systems. Usually, a DNA sequence encoding the N-terminal portion of an endogenous yeast protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the yeast or human superoxide dismutase (SOD) gene, can be linked at the 5' terminus of a foreign gene and

expressed in yeast. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See e.g., EP-A-0 196 056. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g., ubiquitin specific processing protease) to cleave the ubiquitin from the foreign protein. Through this method, therefore, native foreign protein can be isolated (e.g., WO88/024066).

**[0140]** Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provide for secretion in yeast of the foreign protein. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either in vivo or in vitro. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

**[0141]** DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (EP-A-0 012 873; JPO. 62,096,086) and the A-factor gene (US patent 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (EP-A-0 060 057).

**[0142]** A preferred class of secretion leader sequences is that which employs a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (usually about 25 to about 50 amino acid residues) (US Patents 4,546,083 and 4,870,008; EP-A-0 324 274). Additional leaders employing an alpha- factor leader fragment that provides for secretion include hybrid alpha- factor leaders made with a presequence of a first yeast, but a pro-region from a second yeast alpha factor. (e.g., see W 0 89/02463.) Usually, transcription termination sequences recognized by yeast are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes.

**[0143]** Usually, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a prokaryotic host for cloning and amplification, Examples of such yeast-bacteria shuttle vectors include YEp24 (Botstein et al. (1979) Gene 8:17-24), pCl/1 (Brake et al. (1984) PNAS USA 81:4642-4646), and YRp17 (Stinchcomb et al. (1982) J. Mol. Biol. 158:157). In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20. Either a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host. See e.g., Brake et al., supra.

**[0144]** Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homologous sequences flanking the expression construct. Integrations appear to result from recombinations between homologous DNA in the vector and the yeast chromosome (Orr-Weaver et al. (1983) Methods in Enzymol. 101:228-245). An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver et al., supra. One or more expression construct may integrate, possibly affecting levels of recombinant protein produced (Rine et al. (1983) Proc. Natl. Acad. Sci. USA 80:6750). The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or two segments homologous to adjacent segments in the chromosome and flanking the expression construct in the vector, which can result in the stable integration of only the expression construct.

**[0145]** Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of yeast strains that have been transformed. Selectable markers may include biosynthetic genes that can be expressed in the yeast host,

such as ADE2, HIS4, LEU2, TRPI, and ALG7, and the G418 resistance gene, which confer resistance in yeast cells to tunicamycin and G418, respectively. In addition, a suitable selectable marker may also provide yeast with the ability to grow in the presence of toxic compounds, such as metal. For example, the presence of CUP1; allows yeast to grow in the presence of copper ions (Butt et al. (1987) Microbiol, Rev. 51:351), Alternatively, some of the above described components can be put together into transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

**[0146]** Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have been developed for transformation into many yeasts. For example, expression vectors have been developed for, inter alia, the following yeasts: *Candida albicans* (Kurtz, et al. (1986) Mol. Cell. Biol. 6:142), *Candida maltosa* (Kunze, et al. (1985) J. Basic Microbiol. 25:141), *Hansenula polymorpha* (Gleeson, et al. (1986) J. Gen. Microbiol. 132:3459; Roggenkamp et al. (1986) Mol. Gen. Genet. 202:302), *Kluyveromyces fragilis* (Das, et al. (1984) J. Bacteriol. 158:1165), *Kluyveromyces lactis* (De Louvencourt et al. (1983) J. Bacteriol. 154:737; Van den Berg et al. (1990) BiolTechnology 8:135), *Pichia guillerimondii* (Kunze et al. (1985) J. Basic Microbiol. 25:141), *Pichia pastoris* (Cregg, et al. (1985) Mol. Cell. Biol. 5:3376; US Patent Nos. 4,837,148 and 4,929, 555), *Saccharomyces cerevisiae* (Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75:1929; Ito et al. (1983) J. Bacteriol. 153:163), *Schizosaccharomyces pombe* (Beach and Nurse (1981) Nature 300:706), and *Yarrowia lipolytica* (Davidow, et al. (1985) Curr. Genet. 10:39; Gaillardin, et al. (1985) Curr. Genet. 10:49).

**[0147]** Methods of introducing exogenous DNA into yeast hosts are well- known in the art, and usually include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. See e.g., (Kurtz et al. (1986) Mol. Cell. Biol. 6:142; Kunze et al. (1985) J. Basic Microbiol. 25:141; Candida); (Gleeson et al. (1986) J. Gen. Microbiol. 132:3459; Roggenkamp et al. (1986) Mol. Gen. Genet. 202:302; Hansenula); (Das et al. (1984) J. Bacteriol. 158:1165; De Louvencourt et al. (1983) J. Bacteriol. 154:1165; Van den Berg et al. (1990) BiolTechnology 8:135; Kluyveromyces); (Cregg et al. (1985) Mol. Cell. Biol. 5:3376; Kunze et al, (1985) J. Basic Microbiol. 25:141; US Patent Nos. 4,837,148 and 4,929,555; Pichia); (Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75;1929; Ito et al. (1983) J.

Bacteriol. 153:163; Saccharomyces); (Beach and Nurse (1981) Nature 300:706; Schizosaccharomyces); and (Davidow et al. (1985) Curr. Genet. 10:39; Gaillardin et al. (1985) Curr. Genet. 10:49; Yarrowia).

Screens

**[0148]** Another aspect of the present invention includes screening of the immunogenic compositions. Such screening may be performed for a wide range of purposes including by way of example selecting more antigenic immunogenic polypeptides to maximize the immune response in the vaccine recipient, screening multi-component vaccine candidates for immune response to all of the components, screening immunogenic polypeptides for no or only limited side effects, and screening for any other characteristic one of skill in the art may desire, non-limiting examples of which may be found throughout the specification.

**[0149]** The immunogenicity of immunogenic polypeptides may be assayed by any method known to one skilled in the art. Typically, the presence (or absence), titres, affinities, avitidies, etc. of antibodies generated in vivo are tested by standard methods, such as, but not limited to, ELISA assays, by which the immunogenicity or antigenicity are tested on immunoglobulin present in the serum of an organism (or patient). Additional methods, such as generating T-cell hybridomas and measuring activation in the presence of antigen presenting cells ("APCs") and antigen (Surman S et al., 2001 Proc. Natl. Acad. Sci. USA 98: 4587-92, below), examining labeled or unlabeled MHC presented peptides by chromatography, electorphoresis, and/or mass spectroscopy, T-cell activation assays, such as, but not limited to, T cell proliferation assays (Adorini L et al., 1988. J. Exp. Med. 168: 2091; So T. et al., 1996. Immunol. Let. 49: 91-97) and IL-2 production by proliferative response assays of CTLL-2 cells (Gillis S et al., 1978. J. Immunol. 120: 2027; So T. et al., 1996. Immunol. Let. 49: 91-97), and many others may be applied to determine more specific aspects of an immune response, or the lack thereof, such as, for example, the identity of the immunogenic T cell epitope of the antigen.

**[0150]** As non-limiting, specific examples, in vitro T cell assays may be carried out whereby the polypeptide, protein, or protein complex can be processed and presented in the groove of MHC molecules by appropriate antigen-presenting cells (APCs) to syngeneic T cells. T cell responses may be measured by simple proliferation measurements or by measuring release of specific cytokine by activated cells; APCs may be irradiated or otherwise treated to prevent

proliferation to facilitate interpretation of the results of such assays. In order to determine the immunogenicity of an epitope in the context of different MHC allotypes, in vivo assays using syngeneic APCs and T-cells of a range of allotypes may be carried out to test for T cell epitopes in a range of individuals or patients.

**[0151]** Alternatively, transgenic animals expressing MHC molecules from human (or any other species of interest) maybe used to assay for T cell epitopes; in a preferred embodiment this assay is carried out in transgenic animals in which the endogenous MHC repertoire has been knocked out and, better yet, in which one or more other accessory molecules of the endogenous MHC/T cell receptor complex have also been replaced with human molecules (or molecules of any other species of interest), such as, for example, the CD4 molecule.

**[0152]** Furthermore, to detect anti-protein/antigen/immunogenic polypeptide antibodies directly in vivo, for example in clinical and animal studies, ELISA assays, such as, for example solid phase indirect ELISA assays, may be used to detect binding of antibodies. In one specific embodiment, microtiter plates are incubated with the immunogenic polypeptide of interest at an appropriate concentration and in a suitable buffer. After washes with an appropriate washing solution, such as, for example PBS (pH 7.4), PBS containing 1% BSA and 0.05% Tween 20, or any other such solution as may be appropriate, serum samples are diluted, for example in PBS/BSA, and equal volumes of the samples are added in duplicate to the wells. The plates are incubated, and after additional washes, for example with PBS, anti-immunoglobulin antibodies coupled/conjugated to a reporter, such as a radioactive isotope or alkaline phosphatase, are added to each well at an appropriate concentration, and incubated. The wells are then washed again, and for example, in the case of use of alkaline phosphatase as a reporter, the enzyme reaction is carried our using a colorometric substrate, such as p-nitrophenyl phosphate in diethanolamine buffer (pH 9.8), absorbance of which can be read at 405 nm, for example, in an automatic ELISA reader (e.g. Multiskan PLUS; Labsystems).

**[0153]** As an additional non-limiting example, to detect antibodies in the serum of patients and animals, immunoblotting can also be applied. In one specific embodiment, an appropriate amount of the immunogenic polypeptide of interest per samples/lane is run on gels (e.g. polyacrylamide), under reducing and/or nonreducing conditions, and the polypeptide is transferred to a membrane, such as, for example, PVDF membranes; any other method to separate proteins by size can be used followed by transfer of the polypeptide to a membrane. The membranes are blocked, for example, using a solution of 5% (w/v) milk powder in PBS.

In another embodiment, purified immunogenic polypeptide may be applied to the membrane. The blots are then incubated with serum samples at varying dilutions in the blocking solution (before and after injection regimen) and control anti-antigen, so far as such samples are available. The blots will be washed four times with an appropriate washing solution, and further incubated with reporter-conjugated anti-immunoglobulin at a appropriate/specified dilutions for appropriate/specified periods of time under appropriate/specified conditions. The blots are washed again with an appropriate washing solution, and the immunoreactive protein bands are visualized, for example, in the case of use of horseradish peroxidase-conjugated anti-immunoglobulin, using enhanced chemiluminescence reagents marketed by Amersham (Bucks, United Kingdom).

**[0154]** To test for a neutralizing effect of antibodies generated in vivo (patients or animals), a relevant biological activity of the pathogen of interest can, for example, be determined by using the bioassays, such, as for example, cell proliferation assays or host adhesion, in varying concentrations of serum of individuals or animals exposed/immunized with the immunogenic polypeptide of interest. Exponentially growing cells of the pathogen are washed and resuspended to a consistent and appropriate concentration in growth medium in a series of serial dilutions, and added in aliquots to each well. For neutralization, a dilution series of serum before and after in vivo exposure (immunization) is added to the wells. The plates are incubated for an appropriate period of time (depending on the pathogen). The growth rate of the pathogen in each well is determined.

**[0155]** A preferred method of screening for immunogenicity is by the Active Maternal Immunization Assay. As discussed in Example 1, this assay may be used to measure serum titers of the female mice during the immunization schedule as well as the survival time of the pups after challenge. The skilled artisan can use the other methods of screening to determine antigenicity or immunogenicity of the immunogenic polypeptides of the present invention set forth in this specification and in the art for screening immunogenic polypeptides.

**[0156]** Methods of screening for antigenicity or immunogenicity may be used to select immunogenic polypeptides of interest from groups of two or more, three or more, five or more, ten or more, or fifty or more immunogenic polypeptides of the present invention based upon a criterion. One of skill in the art may apply any desired criterion in selecting the immunogenic polypeptide of interest. The criterion will depend upon the intended use of the immunogenic polypeptide of interest. By way of example, but not limitation, the criterion

may be as simple as selecting the polypeptide with the highest antigenicity or immunogenicity. More complicated criterion may also be used such as selecting the polypeptide with the highest antigenicity or immunogenicity that produces no undesirable side effects upon immunization or selecting a multicomponent vaccine that includes the immunogenic polypeptide that has the highest antigenicity or immunogenicity against a panel of pathogens. Determination of the criterion is a simple matter of experimental design based upon the intended use and therefore one of skill in the art would have no difficulty in selecting appropriate criteria for any situation.

Antibodies

**[0157]** As used herein, the term "antibody" refers to a polypeptide or group of polypeptides composed of at least one antibody combining site. An "antibody combining site" is the three-dimensional binding space with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows a binding of the antibody with the antigen. Antibody includes, for example, vertebrate antibodies, hybrid antibodies, chimeric antibodies, humanized antibodies, altered antibodies, univalent antibodies, Fab proteins, and single domain antibodies.

**[0158]** Antibodies against the proteins of the invention are useful for affinity chromatography, immunoassays, and distinguishing/identifying meningococcal proteins.

**[0159]** Antibodies to the proteins of the invention, both polyclonal and monoclonal, may be prepared by conventional methods. In general, the protein is first used to immunize a suitable animal, preferably a mouse, rat, rabbit or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 $\mu$g/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by in vitro immunization using methods known in the art, which for the purposes of this invention is considered equivalent to in vivo immunization. Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25° C

for one hour, followed by incubating at 40° C for 2-18 hours. The serum is recovered by centrifugation (e.g., 1,000 g for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

**[0160]** Monoclonal antibodies are prepared using the standard method of Kohler & Milstein (Nature (1975) 256:495-96), or a modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells may be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate or well coated with the protein antigen, B- cells expressing membrane-bound immunoglobulin specific for the antigen bind to the plate, and are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected MAb- secreting hybridomas are then cultured either in vitro (e.g. , in tissue culture bottles or hollow fiber reactors), or in vivo (as ascites in mice).

**[0161]** If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques. Suitable labels include fluorophores, chromophores, radioactive atom s (particularly $^{32}P$ and $^{125}I$), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TMB) to a blue pigment, quantifiable with a spectrophotometer. "Specific binding partner" refers to a protein capable of binding a ligand molecule with high specificity, as for example in the case of an antigen and a monoclonal antibody specific therefor. Other specific binding partners include biotin and avidin or streptavidin, IgG and protein A, and the numerous receptor-ligand couples known in the art. It should be understood that the above description is not meant to categorize the various labels into distinct classes, as the same label may serve in several different modes. For example, $^{125}I$ may serve as a radioactive label or as an electron- dense reagent. HRP may serve as enzyme or as antigen for a MAb. Further, one may combine various labels for desired effect. For example, MAbs and avidin also require labels in the practice of this invention: thus, one might label a

MAb with biotin, and detect its presence with avidin labeled with $^{125}I$, or with an anti-biotin MAb labeled with HRP. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the invention.

Pharmaceutical Compositions

**[0162]** Pharmaceutical compositions can comprise either polypeptides, antibodies, or nucleic acid of the invention. The pharmaceutical compositions will comprise a therapeutically effective amount of either polypeptides, antibodies, or polynucleotides of the claimed invention.

**[0163]** The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature. The precise effective amount for a subject will depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. Thus, it is not useful to specify an exact effective amount in advance. However, the effective amount for a given situation can be determined by routine experimentation and is within the judgment of the clinician.

**[0164]** For purposes of the present invention, an effective dose will be from about 0.01 mg/ kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered. Preferred dosages for protein based pharmaceuticals including vaccines will be between 5 and 500 $\mu$5 of the immunogenic polypeptides of the present invention.

**[0165]** A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids,

polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

**[0166]** Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

**[0167]** Pharmaceutically acceptable carriers in therapeutic compositions may contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Typically, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier.

Delivery Methods

**[0168]** Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

**[0169]** Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal or transcutancous applications (e. g., see WO98/20734), needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Vaccines

**[0170]** Vaccines according to the invention may either be prophylactic (i.e., to prevent infection) or therapeutic (i.e., to treat disease after infection).

**[0171]** Such vaccines comprise immunogenic polypeptide(s), immunogen(s), polypeptide(s), protein(s) or nucleic acid, usually in combination with "pharmaceutically

acceptable carriers," which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the antigen or immunogen may be conjugated to a bacterial toxoid, such as a toxoid from such pathogens as diphtheria, tetanus, cholera, H. pylori, etc.

**[0172]** Compositions such as vaccines and pharmaceutical compositions of the invention may advantageously include an adjuvant, which can function to enhance the immune responses (humoral and/or cellular) elicited in a patient who receives the composition.

**[0173]** Adjuvants that can be used with the invention include, but are not limited to:

- A mineral containing composition, including calcium salts and aluminum salts (or mixtures thereof). Calcium salts include calcium phosphate (e.g. the "CAP" particles disclosed in ref. 79, which is hereby incorporated by reference for all of its teachings with particular reference to "CAP" particles). Aluminum salts include hydroxides, phosphates, sulfates, etc., with the salts taking any suitable form (e.g. gel, crystalline, amorphous, etc.). Adsorption to these salts is preferred. The mineral containing compositions may also be formulated as a particle of metal salt (Reference 80). Aluminum salt adjuvants are described in more detail below.

- Cytokine inducing agents (see in more detail below).

- Saponins (chapter 22 of ref. 81), which are a heterologous group of sterol glycosides and triterpenoid glycosides that are found in the bark, leaves, stems, roots and even flowers of a wide range of plant species. Saponin from the bark of the Quillaia saponaria Molina tree have been widely studied as adjuvants. Saponin can also be commercially obtained from Smilax ornata (sarsaparilla), Gypsophilla paniculata (brides veil), and Saponaria officianalis (soap root). Saponin adjuvant formulations include purified formulations, such as QS21, as well as lipid formulations, such as ISCOMs. QS21 is marketed as Stimulon™. Saponin compositions have been purified using HPLC and RP-HPLC. Specific purified fractions using these techniques have been identified, including

QS7, QS17, QS18, QS21, QH-A, QH-B and QH-C. Preferably, the saponin is QS21. A method of production of QS21 is disclosed in ref. 82 (which is hereby incorporated by reference for all its teachings with particular references to methods of production and use of QS7, QS17, QS18 and QS21). It is possible to use fraction A of Quil A together with at least one other adjuvant (Ref. 83). Saponin formulations may also comprise a sterol, such as cholesterol (84). Combinations of saponins and cholesterols can be used to form unique particles called immunostimulating complexes (ISCOMs) (chapter 23 of ref. 81). ISCOMs typically also include a phospholipid such as phosphatidylethanolamine or phosphatidylcholine. Any known saponin can be used in ISCOMs. Preferably, the ISCOM includes one or more of QuilA, QHA & QHC. ISCOMs are further described in refs. 85-88 (which is hereby incorporated by reference for all its teachings with particular references to ISCOMs, methods of manufacture of ISCOMs and methods of use of ISCOMs).. Optionally, the ISCOMS may be devoid of additional detergent (Ref. 89). It is possible to use a mixture of at least two ISCOM complexes, each complex comprising essentially one saponin fraction, where the complexes are ISCOM complexes or ISCOM matrix complexes (90). A review of the development of saponin based adjuvants can be found in refs. 91 and 92.

- Fatty adjuvants (see in more detail below).

- Bacterial ADP-ribosylating toxins (e.g. the *E. coli* heat labile enterotoxin "LT", cholera toxin "CT", or pertussis toxin "PT") and detoxified derivatives thereof, such as the mutant toxins known as LT-K63 and LT R72 (93). The use of detoxified ADP-ribosylating toxins as mucosal adjuvants is described in ref. 94 and as parenteral adjuvants in ref. 95.

- Bioadhesives and mucoadhesives, such as esterified hyaluronic acid microspheres (96) or chitosan and its derivatives (97).

- Microparticles (i.e. a particle of ~100nm to ~150μm in diameter, more preferably ~200nm to ~30μm in diameter, or ~500nm to ~10μm in diameter) formed from materials that are biodegradable and non toxic (e.g. a poly(α-hydroxy acid), a polyhydroxybutyric acid, a polyorthoester, a polyanhydride, a polycaprolactone, etc.), with poly(lactide co glycolide) being preferred, optionally treated to have a negatively-charged surface (e.g. with SDS) or a positively-charged surface (e.g. with a cationic detergent, such as CTAB).

- Liposomes (Chapters 13 & 14 of ref. 81). Examples of liposome formulations suitable for use as adjuvants are described in refs. 98 - 100.

- Oil in water emulsions (see in more detail below).

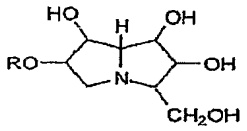
- Polyoxyethylene ethers and polyoxyethylene esters (101). Such formulations further include polyoxyethylene sorbitan ester surfactants in combination with an octoxynol (102) as well as polyoxyethylene alkyl ethers or ester surfactants in combination with at least one additional non-ionic surfactant such as an octoxynol (103). Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether (laureth 9), polyoxyethylene-9-steoryl ether, polyoxytheylene-8-steoryl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether.

- Muramyl peptides, such as N-acetylmuramyl-L-threonyl-D-isoglutamine ("thr-MDP"), N acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylglucsaminyl-N-acetylmuramyl-L-Al-D-isoglu-L-Ala-dipalmitoxy propylamide ("DTP-DPP", or "Theramide™), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine ("MTP-PE").

- An outer membrane protein proteosome preparation prepared from a first Gram-negative bacterium in combination with a liposaccharide preparation derived from a second Gram negative bacterium, wherein the outer membrane protein proteosome and liposaccharide preparations form a stable non-covalent adjuvant complex. Such complexes include "IVX-908", a complex comprised of Neisseria meningitidis outer membrane and lipopolysaccharides. They have been used as adjuvants for influenza vaccines (104).

- Methyl inosine 5'-monophosphate ("MIMP") (105).

- A polyhydroxlated pyrrolizidine compound (106), such as one having formula:



where R is selected from the group comprising hydrogen, straight or branched, unsubstituted or substituted, saturated or unsaturated acyl, alkyl (e.g. cycloalkyl), alkenyl, alkynyl and aryl groups, or a pharmaceutically acceptable salt or derivative thereof. Examples include, but are not limited to: casuarine, casuarine-6-α-D-glucopyranose, 3 epi casuarine, 7 epi casuarine, 3,7 diepi casuarine, etc.

- A gamma inulin (107) or derivative thereof, such as algammulin.
- A CD1d ligand, such as a α glycosylceramide e.g. α-galactosylceramide.

**[0174]** These and other adjuvant active substances are discussed in more detail in references 81 & 108.

**[0175]** Compositions may include two or more of said adjuvants. For example, they may advantageously include both an oil in water emulsion and a cytokine inducing agent, as this combination improves the cytokine responses elicited by influenza vaccines, such as the interferon γ response, with the improvement being much greater than seen when either the emulsion or the agent is used on its own.

**[0176]** Antigens and adjuvants in a composition will typically be in admixture.

**[0177]** Where a vaccine includes an adjuvant, it may be prepared extemporaneously, at the time of delivery. Thus the invention provides kits including the antigen and adjuvant components ready for mixing. The kit allows the adjuvant and the antigen to be kept separately until the time of use. The components are physically separate from each other within the kit, and this separation can be achieved in various ways. For instance, the two components may be in two separate containers, such as vials. The contents of the two vials can then be mixed e.g. by removing the contents of one vial and adding them to the other vial, or by separately removing the contents of both vials and mixing them in a third container. In a preferred arrangement, one of the kit components is in a syringe and the other is in a container such as a vial. The syringe can be used (e.g. with a needle) to insert its contents into the second container for mixing, and the mixture can then be withdrawn into the syringe. The mixed contents of the syringe can then be administered to a patient, typically through a new sterile needle. Packing one component in a syringe eliminates the need for using a separate syringe for patient administration. In another preferred arrangement, the two kit components are held together but separately in the same syringe e.g. a dual chamber

syringe, such as those disclosed in references 109-116 etc. When the syringe is actuated (e.g. during administration to a patient) then the contents of the two chambers are mixed. This arrangement avoids the need for a separate mixing step at the time of use.

*Oil in water emulsion adjuvants*

**[0178]** Oil in water emulsions have been found to be particularly suitable for use in adjuvanting influenza virus vaccines. Various such emulsions are known, and they typically include at least one oil and at least one surfactant, with the oil(s) and surfactant(s) being biodegradable (metabolisable) and biocompatible. The oil droplets in the emulsion are generally less than 5μm in diameter, and may even have a sub micron diameter, with these small sizes being achieved with a microfluidiser to provide stable emulsions. Droplets with a size less than 220 nm are preferred as they can be subjected to filter sterilization.

**[0179]** The invention can be used with oils such as those from an animal (such as fish) or vegetable source. Sources for vegetable oils include nuts, seeds and grains. Peanut oil, soybean oil, coconut oil, and olive oil, the most commonly available, exemplify the nut oils. Jojoba oil can be used e.g. obtained from the jojoba bean. Seed oils include safflower oil, cottonseed oil, sunflower seed oil, sesame seed oil and the like. In the grain group, corn oil is the most readily available, but the oil of other cereal grains such as wheat, oats, rye, rice, teff, triticale and the like may also be used. 6-10 carbon fatty acid esters of glycerol and 1,2-propanediol, while not occurring naturally in seed oils, may be prepared by hydrolysis, separation and esterification of the appropriate materials starting from the nut and seed oils. Fats and oils from mammalian milk are metabolizable and may therefore be used in the practice of this invention. The procedures for separation, purification, saponification and other means necessary for obtaining pure oils from animal sources are well known in the art. Most fish contain metabolizable oils which may be readily recovered. For example, cod liver oil, shark liver oils, and whale oil such as spermaceti exemplify several of the fish oils which may be used herein. A number of branched chain oils are synthesized biochemically in 5-carbon isoprene units and are generally referred to as terpenoids. Shark liver oil contains a branched, unsaturated terpenoids known as squalene, 2,6,10,15,19,23-hexamethyl-2,6,10,14,18,22-tetracosahexaene, which is particularly preferred herein. Squalane, the saturated analog to squalene, is also a preferred oil. Fish oils, including squalene and squalane, are readily available from commercial sources or may be obtained by methods

known in the art. Other preferred oils are the tocopherols (see below). Mixtures of oils can be used.

**[0180]** Surfactants can be classified by their 'HLB' (hydrophile/lipophile balance). Preferred surfactants of the invention have a HLB of at least 10, preferably at least 15, and more preferably at least 16. The invention can be used with surfactants including, but not limited to: the polyoxyethylene sorbitan esters surfactants (commonly referred to as the Tweens), especially polysorbate 20 and polysorbate 80; copolymers of ethylene oxide (EO), propylene oxide (PO), and/or butylene oxide (BO), sold under the DOWFAX™ tradename, such as linear EO/PO block copolymers; octoxynols, which can vary in the number of repeating ethoxy (oxy-1,2-ethanediyl) groups, with octoxynol-9 (Triton X 100, or t octylphenoxypolyethoxyethanol) being of particular interest; (octylphenoxy)polyethoxyethanol (IGEPAL CA-630/NP-40); phospholipids such as phosphatidylcholine (lecithin); polyoxyethylene fatty ethers derived from lauryl, cetyl, stearyl and oleyl alcohols (known as Brij surfactants), such as triethyleneglycol monolauryl ether (Brij 30); and sorbitan esters (commonly known as the SPANs), such as sorbitan trioleate (Span 85) and sorbitan monolaurate. Preferred surfactants for including in the emulsion are Tween 80 (polyoxyethylene sorbitan monooleate), Span 85 (sorbitan trioleate), lecithin and Triton X 100. Mixtures of surfactants can be used e.g. Tween 80/Span 85 mixtures.

**[0181]** Specific oil in water emulsion adjuvants useful with the invention include, but are not limited to:

- A submicron emulsion of squalene, Tween 80, and Span 85. The composition of the emulsion by volume can be about 5% squalene, about 0.5% polysorbate 80 and about 0.5% Span 85. In weight terms, these ratios become 4.3% squalene, 0.5% polysorbate 80 and 0.48% Span 85. This adjuvant is known as 'MF59' (117-119), as described in more detail in Chapter 10 of ref. 81 and chapter 12 of ref. 108. The MF59 emulsion advantageously includes citrate ions e.g. 10mM sodium citrate buffer.

- An emulsion of squalene, a tocopherol, and Tween 80. The emulsion may include phosphate buffered saline. It may also include Span 85 (e.g. at 1%) and/or lecithin. These emulsions may have from 2 to 10% squalene, from 2 to 10% tocopherol and from 0.3 to 3% Tween 80, and the weight ratio of squalene:tocopherol is preferably <1 as this provides a more stable emulsion. One such emulsion can be made by dissolving Tween

80 in PBS to give a 2% solution, then mixing 90ml of this solution with a mixture of (5g of DL α tocopherol and 5ml squalene), then microfluidising the mixture. The resulting emulsion may have submicron oil droplets e.g. with an average diameter of between 100 and 250nm, preferably about 180nm.

- An emulsion of squalene, a tocopherol, and a Triton detergent (e.g. Triton X-100).

- An emulsion of squalane, polysorbate 80 and poloxamer 401 ("Pluronic™ L121"). The emulsion can be formulated in phosphate buffered saline, pH 7.4. This emulsion is a useful delivery vehicle for muramyl dipeptides, and has been used with threonyl MDP in the "SAF 1" adjuvant (120) (0.05-1% Thr MDP, 5% squalane, 2.5% Pluronic L121 and 0.2% polysorbate 80). It can also be used without the Thr MDP, as in the "AF" adjuvant (121) (5% squalane, 1.25% Pluronic L121 and 0.2% polysorbate 80). Microfluidisation is preferred.

- An emulsion having from 0.5 50% of an oil, 0.1 10% of a phospholipid, and 0.05 5% of a non ionic surfactant. As described in reference 122, preferred phospholipid components are phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, phosphatidylinositol, phosphatidylglycerol, phosphatidic acid, sphingomyelin and cardiolipin. Submicron droplet sizes are advantageous.

- A submicron oil-in-water emulsion of a non-metabolisable oil (such as light mineral oil) and at least one surfactant (such as lecithin, Tween 80 or Span 80). Additives may be included, such as QuilA saponin, cholesterol, a saponin-lipophile conjugate (such as GPI-0100, described in reference 123, produced by addition of aliphatic amine to desacylsaponin via the carboxyl group of glucuronic acid), dimethyidioctadecylammonium bromide and/or N,N-dioctadecyl-N,N-bis (2-hydroxyethyl)propanediamine.

- An emulsion in which a saponin (e.g. QuilA or QS21) and a sterol (e.g. a cholesterol) are associated as helical micelles (124).

**[0182]** The emulsions may be mixed with antigen extemporaneously, at the time of delivery. Thus the adjuvant and antigen may be kept separately in a packaged or distributed vaccine, ready for final formulation at the time of use. The antigen will generally be in an aqueous

form, such that the vaccine is finally prepared by mixing two liquids. The volume ratio of the two liquids for mixing can vary (e.g. between 5:1 and 1:5) but is generally about 1:1.

**[0183]** After the antigen and adjuvant have been mixed, the antigen will generally remain in aqueous solution but may distribute itself around the oil/water interface. In general, little if any antigen will enter the oil phase of the emulsion.

**[0184]** Where a composition includes a tocopherol, any of the α, β, γ, δ, ε or ζ tocopherols can be used, but α tocopherols are preferred. The tocopherol can take several forms e.g. different salts and/or isomers. Salts include organic salts, such as succinate, acetate, nicotinate, etc. D α tocopherol and DL α tocopherol can both be used. Tocopherols are advantageously included in vaccines for use in elderly patients (e.g. aged 60 years or older) because vitamin E has been reported to have a positive effect on the immune response in this patient group (125). They also have antioxidant properties that may help to stabilize the emulsions (126). A preferred α tocopherol is DL α tocopherol, and the preferred salt of this tocopherol is the succinate. The succinate salt has been found to cooperate with TNF related ligands in vivo. Moreover, α tocopherol succinate is known to be compatible with vaccines (for example, influenza vaccines) and to be a useful preservative as an alternative to mercurial compounds (127).

**Cytokine-inducing agents**

**[0185]** Cytokine inducing agents for inclusion in compositions of the invention are able, when administered to a patient, to elicit the immune system to release cytokines, including interferons and interleukins. Cytokine responses are known to be involved in the early and decisive stages of host defense against pathogen infection (128). Preferred agents can elicit the release of one or more of: interferon γ; interleukin 1; interleukin 2; interleukin 12; TNF α; TNF β; and GM CSF. Preferred agents elicit the release of cytokines associated with a Th1-type immune response e.g. interferon γ, TNF α, interleukin 2. Stimulation of both interferon γ and interleukin 2 is preferred.

**[0186]** As a result of receiving a composition of the invention, therefore, a patient will have T cells that, when stimulated with an antigen, will release the desired cytokine(s) in an antigen specific manner. For example, T cells purified form their blood will release γ interferon when exposed in vitro to the stimulated antigen. Methods for measuring such responses in peripheral blood mononuclear cells (PBMC) are known in the art, and include

ELISA, ELISPOT, flow cytometry and real time PCR. For example, reference 129 reports a study in which antigen specific T cell-mediated immune responses against tetanus toxoid, specifically γ interferon responses, were monitored, and found that ELISPOT was the most sensitive method to discriminate antigen specific TT-induced responses from spontaneous responses, but that intracytoplasmic cytokine detection by flow cytometry was the most efficient method to detect re stimulating effects.

**[0187]** Suitable cytokine inducing agents include, but are not limited to:

- An immunostimulatory oligonucleotide, such as one containing a CpG motif (a dinucleotide sequence containing an unmethylated cytosine linked by a phosphate bond to a guanosine), or a double stranded RNA, or an oligonucleotide containing a palindromic sequence, or an oligonucleotide containing a poly(dG) sequence.

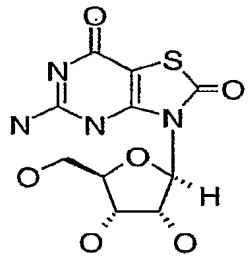
- 3 O deacylated monophosphoryl lipid A ('3dMPL', also known as 'MPL™') (130-133).

- An imidazoquinoline compound, such as Imiquimod ("R 837") (134, 135), Resiquimod ("R 848") (136), and their analogs; and salts thereof (e.g. the hydrochloride salts). Further details about immunostimulatory imidazoquinolines can be found in references 137 to 141.

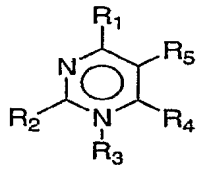
- A thiosemicarbazone compound, such as those disclosed in reference 142. Methods of formulating, manufacturing, and screening for active compounds are also described in reference 142. The thiosemicarbazones are particularly effective in the stimulation of human peripheral blood mononuclear cells for the production of cytokines, such as TNF-α.

- A tryptanthrin compound, such as those disclosed in reference 143. Methods of formulating, manufacturing, and screening for active compounds are also described in reference 143. The thiosemicarbazones are particularly effective in the stimulation of human peripheral blood mononuclear cells for the production of cytokines, such as TNF-α.

- A nucleoside analog, such as: (a) Isatorabine (ANA-245; 7-thia-8-oxoguanosine):



**[0188]** and prodrugs thereof; (b) ANA975; (c) ANA-025-1; (d) ANA380; (e) the compounds disclosed in references 144 to 146; (f) a compound having the formula:

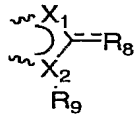


**[0189]** wherein:

R1 and R2 are each independently H, halo, -NRaRb, -OH, C1-6 alkoxy, substituted C1-6 alkoxy, heterocyclyl, substituted heterocyclyl, C6-10 aryl, substituted C6-10 aryl, C1-6 alkyl, or substituted C1-6 alkyl;

R3 is absent, H, C1-6 alkyl, substituted C1-6 alkyl, C6-10 aryl, substituted C6-10 aryl, heterocyclyl, or substituted heterocyclyl;

R4 and R5 are each independently H, halo, heterocyclyl, substituted heterocyclyl, C(O)-Rd, C1-6 alkyl, substituted C1-6 alkyl, or bound together to form a 5 membered ring as in R4-5:



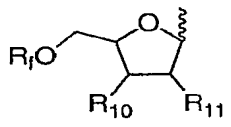
$R_{4-5}$

the binding being achieved at the bonds indicated by a ~~~

X1 and X2 are each independently N, C, O, or S;

R8 is H, halo, -OH, C1-6 alkyl, C2-6 alkenyl, C2-6 alkynyl, -OH, -NRaRb, -(CH2)n-O-Rc, -O-(C1-6 alkyl), -S(O)pRe, or -C(O)-Rd;

R9 is H, C1-6 alkyl, substituted C1-6 alkyl, heterocyclyl, substituted heterocyclyl or R9a, wherein R9a is:



R9a

the binding being achieved at the bond indicated by a ~~~

R10 and R11 are each independently H, halo, C1-6 alkoxy, substituted C1-6 alkoxy, -NRaRb, or -OH;

each Ra and Rb is independently H, C1-6 alkyl, substituted C1-6 alkyl, -C(O)Rd, C6-10 aryl;

each Rc is independently H, phosphate, diphosphate, triphosphate, C1-6 alkyl, or substituted C1-6 alkyl;

each Rd is independently H, halo, C1-6 alkyl, substituted C1-6 alkyl, C1-6 alkoxy, substituted C1-6 alkoxy, -NH2, -NH(C1-6 alkyl), -NH(substituted C1-6 alkyl), -N(C1-6 alkyl)2, -N(substituted C1-6 alkyl)2, C6-10 aryl, or heterocyclyl;

each Re is independently H, C1-6 alkyl, substituted C1-6 alkyl, C6-10 aryl, substituted C6-10 aryl, heterocyclyl, or substituted heterocyclyl;

each Rf is independently H, C1-6 alkyl, substituted C1-6 alkyl, -C(O)Rd, phosphate, diphosphate, or triphosphate;

each n is independently 0, 1, 2, or 3;

each p is independently 0, 1, or 2; or

or (g) a pharmaceutically acceptable salt of any of (a) to (f), a tautomer of any of (a) to (f), or a pharmaceutically acceptable salt of the tautomer.

- Loxoribine (7-allyl-8-oxoguanosine) (147).

- Compounds disclosed in reference 148, including: Acylpiperazine compounds, Indoledione compounds, Tetrahydraisoquinoline (THIQ) compounds, Benzocyclodione compounds, Aminoazavinyl compounds, Aminobenzimidazole quinolinone (ABIQ) compounds (149, 150), Hydrapthalamide compounds, Benzophenone compounds, Isoxazole compounds, Sterol compounds, Quinazilinone compounds, Pyrrole compounds

(151), Anthraquinone compounds, Quinoxaline compounds, Triazine compounds, Pyrazalopyrimidine compounds, and Benzazole compounds (152).

- Compounds disclosed in reference 153.
- An aminoalkyl glucosaminide phosphate derivative, such as RC 529 (154, 155).
- A phosphazene, such as poly(di(carboxylatophenoxy)phosphazene) ("PCPP") as described, for example, in references 156 and 157.
- Small molecule immunopotentiators (SMIPs) such as:

N2-methyl-1-(2-methylpropyl)-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2,N2-dimethyl-1-(2-methylpropyl)-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2-ethyl-N2-methyl-1-(2-methylpropyl)-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2-methyl-1-(2-methylpropyl)-N2-propyl-1H-imidazo(4,5-c)quinoline-2,4-diamine

1-(2-methylpropyl)-N2-propyl-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2-butyl-1-(2-methylpropyl)-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2-butyl-N2-methyl-1-(2-methylpropyl)-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2-methyl-1-(2-methylpropyl)-N2-pentyl-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2-methyl-1-(2-methylpropyl)-N2-prop-2-enyl-1H-imidazo(4,5-c)quinoline-2,4-diamine

1-(2-methylpropyl)-2-((phenylmethyl)thio)-1H-imidazo(4,5-c)quinolin-4-amine

1-(2-methylpropyl)-2-(propylthio)-1H-imidazo(4,5-c)quinolin-4-amine

2-((4-amino-1-(2-methylpropyl)-1H-imidazo(4,5-c)quinolin-2-yl)(methyl)amino)ethanol

2-((4-amino-1-(2-methylpropyl)-1H-imidazo(4,5-c)quinolin-2-yl)(methyl)amino)ethyl acetate

4-amino-1-(2-methylpropyl)-1,3-dihydro-2H-imidazo(4,5-c)quinolin-2-one

N2-butyl-1-(2-methylpropyl)-N4,N4-bis(phenylmethyl)-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2-butyl-N2-methyl-1-(2-methylpropyl)-N4,N4-bis(phenylmethyl)-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2-methyl-1-(2-methylpropyl)-N4,N4-bis(phenylmethyl)-1H-imidazo(4,5-c)quinoline-2,4-diamine

N2,N2-dimethyl-1-(2-methylpropyl)-N4,N4-bis(phenylmethyl)-1H-imidazo(4,5-c)quinoline-2,4-diamine

1-{4-amino-2-(methyl(propyl)amino)-1H-imidazo(4,5-c)quinolin-1-yl}-2-methylpropan-2-ol

1-(4-amino-2-(propylamino)-1H-imidazo(4,5-c)quinolin-1-yl)-2-methylpropan-2-ol

N4,N4-dibenzyl-1-(2-methoxy-2-methylpropyl)-N2-propyl-1H-imidazo(4,5-c)quinoline-2,4-diamine.

**[0190]** The cytokine inducing agents for use in the present invention may be modulators and/or agonists of Toll-Like Receptors (TLR). For example, they may be agonists of one or more of the human TLR1, TLR2, TLR3, TLR4, TLR7, TLR8, and/or TLR9 proteins. Preferred agents are agonists of TLR7 (e.g. imidazoquinolines) and/or TLR9 (e.g. CpG oligonucleotides). These agents are useful for activating innate immunity pathways.

**[0191]** The cytokine inducing agent can be added to the composition at various stages during its production. For example, it may be within an antigen composition, and this mixture can then be added to an oil in water emulsion. As an alternative, it may be within an oil in water emulsion, in which case the agent can either be added to the emulsion components before emulsification, or it can be added to the emulsion after emulsification. Similarly, the agent may be coacervated within the emulsion droplets. The location and distribution of the cytokine inducing agent within the final composition will depend on its hydrophilic/lipophilic properties e.g. the agent can be located in the aqueous phase, in the oil phase, and/or at the oil water interface.

**[0192]** The cytokine inducing agent can be conjugated to a separate agent, such as an antigen (e.g. CRM197). A general review of conjugation techniques for small molecules is provided in ref. 158. As an alternative, the adjuvants may be non-covalently associated with additional agents, such as by way of hydrophobic or ionic interactions.

**[0193]** Two preferred cytokine inducing agents are (a) immunostimulatory oligonucleotides and (b) 3dMPL.

**[0194]** Immunostimulatory oligonucleotides can include nucleotide modifications/analogs such as phosphorothioate modifications and can be double-stranded or (except for RNA) single-stranded. References 159, 160 and 161 disclose possible analog substitutions e.g. replacement of guanosine with 2'-deoxy-7-deazaguanosine. The adjuvant effect of CpG oligonucleotides is further discussed in refs. 162-167. A CpG sequence may be directed to TLR9, such as the motif GTCGTT or TTCGTT (168). The CpG sequence may be specific for inducing a Th1 immune response, such as a CpG-A ODN (oligodeoxynucleotide), or it may be more specific for inducing a B cell response, such a CpG-B ODN. CpG-A and CpG-B ODNs are discussed in refs. 169-171. Preferably, the CpG is a CpG-A ODN. Preferably, the CpG oligonucleotide is constructed so that the 5' end is accessible for receptor recognition. Optionally, two CpG oligonucleotide sequences may be attached at their 3' ends to form "immunomers". See, for example, references 168 and 172-174. A useful CpG adjuvant is CpG7909, also known as ProMune™ (Coley Pharmaceutical Group, Inc.).

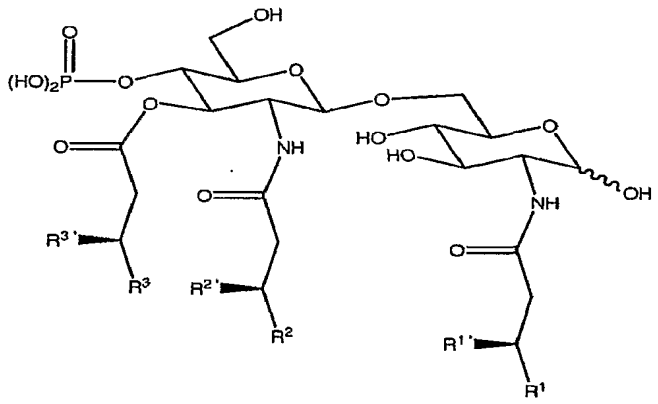
**[0195]** As an alternative, or in addition, to using CpG sequences, TpG sequences can be used (175). These oligonucleotides may be free from unmethylated CpG motifs.

**[0196]** The immunostimulatory oligonucleotide may be pyrimidine rich. For example, it may comprise more than one consecutive thymidine nucleotide (e.g. TTTT, as disclosed in ref. 175), and/or it may have a nucleotide composition with >25% thymidine (e.g. >35%, >40%, >50%, >60%, >80%, etc.). For example, it may comprise more than one consecutive cytosine nucleotide (e.g. CCCC, as disclosed in ref. 148), and/or it may have a nucleotide composition with >25% cytosine (e.g. >35%, >40%, >50%, >60%, >80%, etc.). These oligonucleotides may be free from unmethylated CpG motifs.

**[0197]** Immunostimulatory oligonucleotides will typically comprise at least 20 nucleotides. They may comprise fewer than 100 nucleotides.

**[0198]** 3dMPL (also known as 3 de-O-acylated monophosphoryl lipid A or 3 O desacyl 4' monophosphoryl lipid A) is an adjuvant in which position 3 of the reducing end glucosamine in monophosphoryl lipid A has been de-acylated. 3dMPL has been prepared from a heptoseless mutant of Salmonella minnesota, and is chemically similar to lipid A but lacks an acid-labile phosphoryl group and a base-labile acyl group. It activates cells of the monocyte/macrophage lineage and stimulates release of several cytokines, including IL 1, IL-12, TNF α and GM-CSF (see also ref. 176). Preparation of 3dMPL was originally described in reference 177.

**[0199]** 3dMPL can take the form of a mixture of related molecules, varying by their acylation (e.g. having 3, 4, 5 or 6 acyl chains, which may be of different lengths). The two glucosamine (also known as 2 deoxy-2-amino glucose) monosaccharides are N acylated at their 2 position carbons (i.e. at positions 2 and 2'), and there is also O acylation at the 3' position. The group attached to carbon 2 has formula –NH-CO-CH2-CR1R1'. The group attached to carbon 2' has formula –NH-CO-CH2-CR2R2'. The group attached to carbon 3' has formula –O-CO-CH2-CR3R3'. A representative structure is:



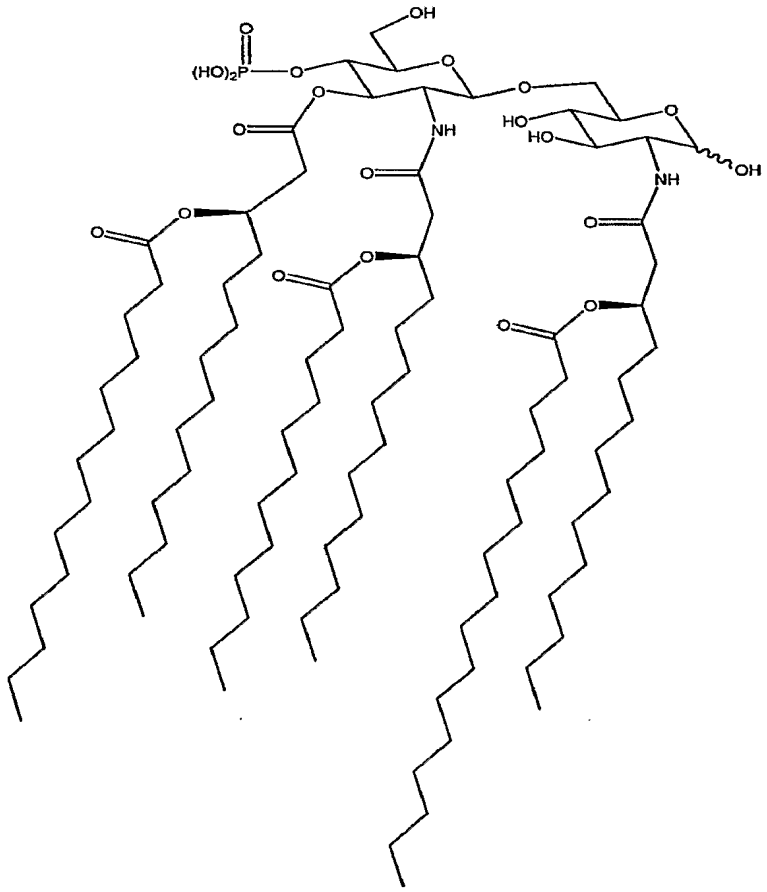

**[0200]** Groups R1, R2 and R3 are each independently –(CH2)n–CH3. The value of n is preferably between 8 and 16, more preferably between 9 and 12, and is most preferably 10.

**[0201]** Groups R1', R2' and R3' can each independently be: (a) –H; (b) –OH; or (c) –O CO R4,where R4 is either –H or –(CH2)m–CH3, wherein the value of m is preferably between 8 and 16, and is more preferably 10, 12 or 14. At the 2 position, m is preferably 14. At the 2' position, m is preferably 10. At the 3' position, m is preferably 12. Groups R1', R2' and R3'

are thus preferably -O acyl groups from dodecanoic acid, tetradecanoic acid or hexadecanoic acid.

**[0202]** When all of R1', R2' and R3' are –H then the 3dMPL has only 3 acyl chains (one on each of positions 2, 2' and 3'). When only two of R1', R2' and R3' are –H then the 3dMPL can have 4 acyl chains. When only one of R1', R2' and R3' is –H then the 3dMPL can have 5 acyl chains. When none of R1', R2' and R3' is –H then the 3dMPL can have 6 acyl chains. The 3dMPL adjuvant used according to the invention can be a mixture of these forms, with from 3 to 6 acyl chains, but it is preferred to include 3dMPL with 6 acyl chains in the mixture, and in particular to ensure that the hexaacyl chain form makes up at least 10% by weight of the total 3dMPL e.g. >20%, >30%, >40%, >50% or more. 3dMPL with 6 acyl chains has been found to be the most adjuvant active form.

**[0203]** Thus the most preferred form of 3dMPL for inclusion in compositions of the invention is:



**[0204]** Where 3dMPL is used in the form of a mixture then references to amounts or concentrations of 3dMPL in compositions of the invention refer to the combined 3dMPL species in the mixture.

**[0205]** In aqueous conditions, 3dMPL can form micellar aggregates or particles with different sizes e.g. with a diameter <150nm or >500nm. Either or both of these can be used with the invention, and the better particles can be selected by routine assay. Smaller particles (e.g. small enough to give a clear aqueous suspension of 3dMPL) are preferred for use according to the invention because of their superior activity (178). Preferred particles have a mean diameter less than 220nm, more preferably less than 200nm or less than 150nm or less than 120nm, and can even have a mean diameter less than 100nm. In most cases, however, the mean diameter will not be lower than 50nm. These particles are small enough to be suitable for filter sterilization. Particle diameter can be assessed by the routine technique of dynamic light scattering, which reveals a mean particle diameter. Where a particle is said to have a diameter of x nm, there will generally be a distribution of particles about this mean, but at least 50% by number (e.g. >60%, >70%, >80%, >90%, or more) of the particles will have a diameter within the range x+25%.

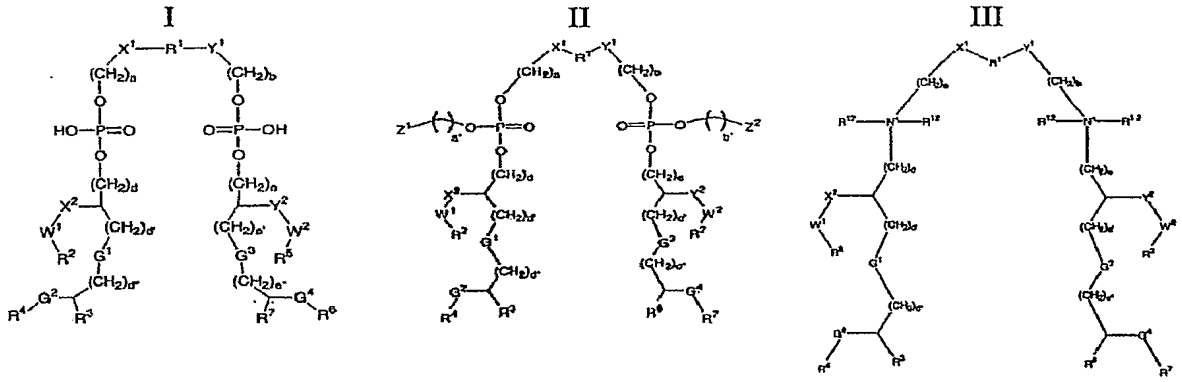
**[0206]** 3dMPL can advantageously be used in combination with an oil in water emulsion. Substantially all of the 3dMPL may be located in the aqueous phase of the emulsion.

**[0207]** The 3dMPL can be used on its own, or in combination with one or more further compounds. For example, it is known to use 3dMPL in combination with the QS21 saponin (179) (including in an oil in water emulsion (180)), with an immunostimulatory oligonucleotide, with both QS21 and an immunostimulatory oligonucleotide, with aluminum phosphate (181), with aluminum hydroxide (182), or with both aluminum phosphate and aluminum hydroxide.

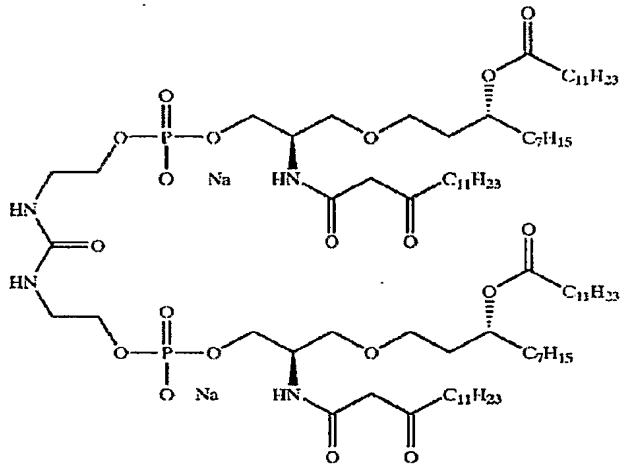
*Fatty adjuvants*

**[0208]** Fatty adjuvants that can be used with the invention include the oil in water emulsions described above, and also include, for example:

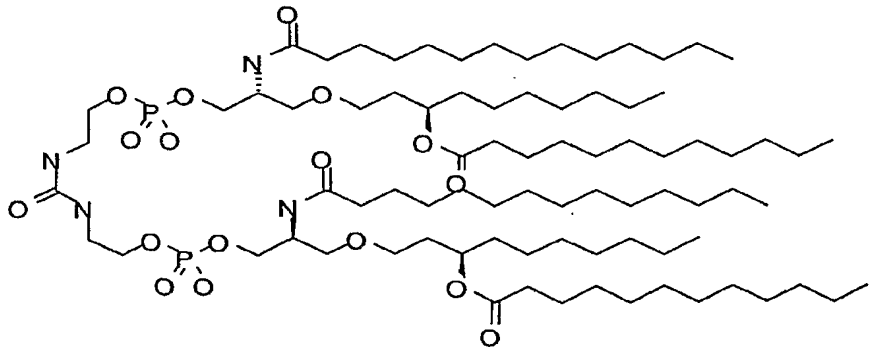
- A compound of formula I, II or III, or a salt thereof:




**[0209]** as defined in reference 183, such as 'ER 803058', 'ER 803732', 'ER 804053', ER 804058', 'ER 804059', 'ER 804442', 'ER 804680', 'ER 804764', ER 803022 or 'ER 804057' e.g.:



ER804057



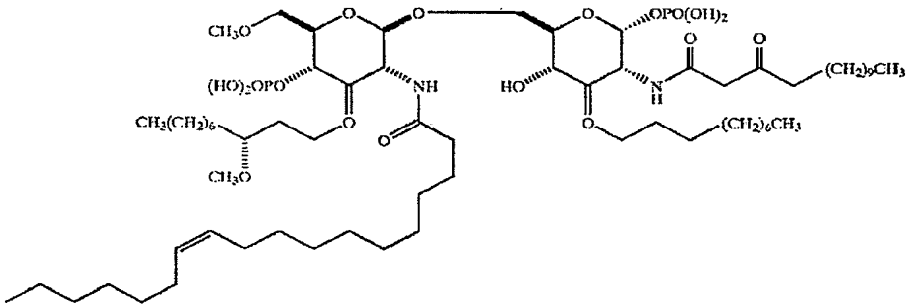
ER 803022:

- Derivatives of lipid A from Escherichia coli such as OM-174 (described in refs. 184 and 185).

- A formulation of a cationic lipid and a (usually neutral) co-lipid, such as aminopropyl-dimethyl-myristoleyloxy-propanaminium bromide-diphytanoylphosphatidyl-ethanolamine ("Vaxfectin™") or aminopropyl-dimethyl-bis-dodecyloxy-propanaminium bromide-dioleoylphosphatidyl-ethanolamine ("GAP-DLRIE:DOPE"). Formulations containing (+)-N-(3-aminopropyl)-N,N-dimethyl-2,3-bis(syn-9-tetradeceneyloxy)-1-propanaminium salts are preferred (186).

- 3 O deacylated monophosphoryl lipid A (see above).

- Compounds containing lipids linked to a phosphate-containing acyclic backbone, such as the TLR4 antagonist E5564 (187, 188):



*Aluminum salt adjuvants*

**[0210]** The adjuvants known as aluminum hydroxide and aluminum phosphate may be used. These names are conventional, but are used for convenience only, as neither is a precise description of the actual chemical compound which is present (e.g. see chapter 9 of reference 81). The invention can use any of the "hydroxide" or "phosphate" adjuvants that are in general use as adjuvants.

**[0211]** The adjuvants known as "aluminium hydroxide" are typically aluminium oxyhydroxide salts, which are usually at least partially crystalline. Aluminium oxyhydroxide, which can be represented by the formula AlO(OH), can be distinguished from other aluminium compounds, such as aluminium hydroxide $Al(OH)_3$, by infrared (IR) spectroscopy, in particular by the presence of an adsorption band at 1070$cm^{-1}$ and a strong shoulder at 3090–3100$cm^{-1}$ (chapter 9 of ref. 81). The degree of crystallinity of an aluminium

hydroxide adjuvant is reflected by the width of the diffraction band at half height (WHH), with poorly crystalline particles showing greater line broadening due to smaller crystallite sizes. The surface area increases as WHH increases, and adjuvants with higher WHH values have been seen to have greater capacity for antigen adsorption. A fibrous morphology (e.g. as seen in transmission electron micrographs) is typical for aluminium hydroxide adjuvants. The pI of aluminium hydroxide adjuvants is typically about 11 i.e. the adjuvant itself has a positive surface charge at physiological pH. Adsorptive capacities of between 1.8-2.6 mg protein per mg $Al^{+++}$ at pH 7.4 have been reported for aluminium hydroxide adjuvants.

**[0212]** The adjuvants known as "aluminium phosphate" are typically aluminium hydroxyphosphates, often also containing a small amount of sulfate (i.e. aluminium hydroxyphosphate sulfate). They may be obtained by precipitation, and the reaction conditions and concentrations during precipitation influence the degree of substitution of phosphate for hydroxyl in the salt. Hydroxyphosphates generally have a $PO_4$/Al molar ratio between 0.3 and 1.2. Hydroxyphosphates can be distinguished from strict $AlPO_4$ by the presence of hydroxyl groups. For example, an IR spectrum band at 3164$cm^{-1}$ (e.g. when heated to 200°C) indicates the presence of structural hydroxyls (ch.9 of ref. 81).

**[0213]** The $PO_4/Al^{3+}$ molar ratio of an aluminium phosphate adjuvant will generally be between 0.3 and 1.2, preferably between 0.8 and 1.2, and more preferably 0.95+0.1. The aluminium phosphate will generally be amorphous, particularly for hydroxyphosphate salts. A typical adjuvant is amorphous aluminium hydroxyphosphate with $PO_4$/Al molar ratio between 0.84 and 0.92, included at 0.6mg $Al^{3+}$/ml. The aluminium phosphate will generally be particulate (e.g. plate like morphology as seen in transmission electron micrographs). Typical diameters of the particles are in the range 0.5-20$\mu$m (e.g. about 5 10$\mu$m) after any antigen adsorption. Adsorptive capacities of between 0.7-1.5 mg protein per mg $Al^{+++}$ at pH 7.4 have been reported for aluminium phosphate adjuvants.

**[0214]** The point of zero charge (PZC) of aluminium phosphate is inversely related to the degree of substitution of phosphate for hydroxyl, and this degree of substitution can vary depending on reaction conditions and concentration of reactants used for preparing the salt by precipitation. PZC is also altered by changing the concentration of free phosphate ions in solution (more phosphate = more acidic PZC) or by adding a buffer such as a histidine buffer (makes PZC more basic). Aluminium phosphates used according to the invention will

generally have a PZC of between 4.0 and 7.0, more preferably between 5.0 and 6.5 e.g. about 5.7.

**[0215]** Suspensions of aluminium salts used to prepare compositions of the invention may contain a buffer (e.g. a phosphate or a histidine or a Tris buffer), but this is not always necessary. The suspensions are preferably sterile and pyrogen free. A suspension may include free aqueous phosphate ions e.g. present at a concentration between 1.0 and 20 mM, preferably between 5 and 15 mM, and more preferably about 10 mM. The suspensions may also comprise sodium chloride.

**[0216]** The invention can use a mixture of both an aluminium hydroxide and an aluminium phosphate (77). In this case there may be more aluminium phosphate than hydroxide e.g. a weight ratio of at least 2:1 e.g. >5:1, >6:1, >7:1, >8:1, >9:1, etc.

**[0217]** The concentration of $Al^{+++}$ in a composition for administration to a patient is preferably less than 10mg/ml e.g. <5 mg/ml, <4 mg/ml, <3 mg/ml, <2 mg/ml, <1 mg/ml, etc. A preferred range is between 0.3 and 1mg/ml.

**[0218]** As well as including one or more aluminium salt adjuvants, the adjuvant component may include one or more further adjuvant or immunostimulating agents. Such additional components include, but are not limited to: a 3-O-deacylated monophosphoryl lipid A adjuvant ('3d MPL'); and/or an oil in water emulsion. 3d MPL has also been referred to as 3 de-O-acylated monophosphoryl lipid A or as 3 O desacyl 4' monophosphoryl lipid A. The name indicates that position 3 of the reducing end glucosamine in monophosphoryl lipid A is de-acylated. It has been prepared from a heptoseless mutant of S.minnesota, and is chemically similar to lipid A but lacks an acid-labile phosphoryl group and a base-labile acyl group. It activates cells of the monocyte/macrophage lineage and stimulates release of several cytokines, including IL-1, IL-12, TNF α and GM-CSF. Preparation of 3d MPL was originally described in reference 150, and the product has been manufactured and sold by Corixa Corporation under the name MPL™. Further details can be found in refs 130 to 133.

**[0219]** Vaccines produced by the invention may be administered to patients at substantially the same time as (e.g. during the same medical consultation or visit to a healthcare professional) other vaccines e.g. at substantially the same time as a measles vaccine, a mumps vaccine, a rubella vaccine, a MMR vaccine, a varicella vaccine, a MMRV vaccine, a diphtheria vaccine, a tetanus vaccine, a pertussis vaccine, a DTP vaccine, a conjugated

H.influenzae type b vaccine, an inactivated poliovirus vaccine, a hepatitis B virus vaccine, a pneumococcal conjugate vaccine, etc. Administration at substantially the same time as a pneumococcal vaccine is particularly useful in elderly patients.

**[0220]** The composition may include an antibiotic.

**[0221]** Immunogenic compositions used as vaccines comprise an immunologically effective amount of the immunogenic polypeptide or immunogenic polypeptides, as well as any other of the above-mentioned components, as needed. By "immunologically effective amount," it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g., nonhuman primate, primate, etc.), the capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

**[0222]** The immunogenic compositions are conventionally administered parenterally, e.g., by injection, either subcutaneously, intramuscularly, or transdermally/transcutaneously (e.g., WO98/20734). Additional formulations suitable for other modes of administration include oral and pulmonary formulations, suppositories, and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents, As an alternative to protein-based vaccines, DNA vaccination may be employed (e.g., Robinson & Torres (1997) Seminars in Immunology 9:271-283; Donnelly et al. (1997) Annu Rev Immunol 15:617-648; see later herein).

Gene Delivery Vehicles

**[0223]** Gene therapy vehicles for delivery of constructs including a coding sequence of a therapeutic of the invention, to be delivered to the mammal for expression in the mammal, can be administered either locally or systemically. These constructs can utilize viral or non-viral vector approaches *in vivo* or *ex vivo*. Expression of such coding sequence can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence *in vivo* can be either constitutive or regulated.

**[0224]** The invention includes gene delivery vehicles capable of expressing the contemplated nucleic acid sequences. The gene delivery vehicle is preferably a viral vector and, more preferably, a retroviral, adenoviral, adeno-associated viral (AAV), herpes viral, or alphavirus vector. The viral vector can also be an astrovirus, coronavirus, orthomyxovirus, papovavirus, paramyxovirus, parvovirus, picornavirus, poxvirus, or togavirus viral vector. See generally, Jolly (1994) Cancer Gene Therapy 1:51-64; Kimura (1994) Human Gene Therapy 5:845-852; Connelly (1995) Human Gene Therapy 6:185-193; and Kaplitt (1994) Nature Genetics 6:148- 153. Retroviral vectors are well known in the art and we contemplate that any retroviral gene therapy vector is employable in the invention, including B, C and D type retroviruses, xenotropic retroviruses (for example, NZB-X1, NZB-X2 and NZB9-1 (see O'Neill (1985) J. Virol. 53:160) polytropic retroviruses e.g., MCF and MCF-MLV (see Kelly (1983) J. Virol. 45:291), spumaviruses and lentiviruses. See RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985.

**[0225]** Portions of the retroviral gene therapy vector may be derived from different retroviruses. For example, retrovector LTRs may be derived from a Murine Sarcoma Virus, a tRNA binding site from a Rous Sarcoma Virus, a packaging signal from a Murine Leukemia Virus, and an origin of second strand synthesis from an Avian Leukosis Virus.

**[0226]** These recombinant retroviral vectors may be used to generate transduction competent retroviral vector particles by introducing them into appropriate packaging cell lines (see US patent 5,591,624). Retrovirus vectors can be constructed for site-specific integration into host cell DNA by incorporation of a chimeric integrase enzyme into the retroviral particle (see WO96/37626). It is preferable that the recombinant viral vector is a replication defective recombinant virus.

**[0227]** Packaging cell lines suitable for use with the above-described retrovirus vectors are well known in the art, are readily prepared (see WO95/30763 and WO92/05266), and can be used to create producer cell lines (also termed vector cell lines or "VCLs") for the production of recombinant vector particles. Preferably, the packaging cell lines are made from human parent cells (e.g., HT1080 cells) or mink parent cell lines, which eliminates inactivation in human serum.

**[0228]** Preferred retroviruses for the construction of retroviral gene therapy vectors include Avian Leukosis Virus, Bovine Leukemia, Virus, Murine Leukemia Virus, Mink-Cell Focus-

Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis Virus and Rous Sarcoma Virus. Particularly preferred Murine Leukemia Viruses include 4070A and 1504A (Hartley and Rowe (1976) J Virol 19:19-25), Abelson (ATCC No. VR-999), Friend (ATCC No. VR-245), Graffi, Gross (ATCC No. VR-590), Kirsten, Harvey Sarcoma Virus and Rauscher (ATCC No. VR-998) and Moloney Murine Leukemia Virus (ATCC No. VR-190). Such retroviruses may be obtained from depositories or collections such as the American Type Culture Collection ("ATCC") in Rockville, Maryland or isolated from known sources using commonly available techniques.

**[0229]** Exemplary known retroviral gene therapy vectors employable in this invention include those described in patent applications GB2200651, EP0415731, EP0345242, EP0334301, WO89/02468; WO89/05349, WO89/09271, WO90/02806, WO90/07936, WO94/03622, WO93/25698, WO93/25234, WO93/11230, WO93/10218, WO91/02805, WO91/02825, WO95/07994, US 5,219,740, US 4,405, 712, US 4,861,719, US 4,980,289, US 4,777,127, US 5,591,624. See also Vile (1993) Cancer Res 53:3860-3864; Vile (1993) Cancer Res 53:962-967; Ram (1993) Cancer Res 53 (1993) 83-88; Takamiya (1992) J Neurosci Res 33:493-503; Baba (1993) J Neurosurg 79:729-735; Mann (1983) Cell 33:153; Cane (1984) Proc Natl Acad Sci 81:6349; and Miller (1990) Human Gene Therapy 1.

**[0230]** Human adenoviral gene therapy vectors are also known in the art and employable in this invention. See, for example, Berkner (1988) Biotechniques 6:616 and Rosenfeld (1991) Science 252:431, and WO93/07283, WO93/06223, and WO93/07282. Exemplary known adenoviral gene therapy vectors employable in this invention include those described in the above referenced documents and in WO94/12649, WO93/03769, WO93/19191, WO94/28938, WO95/11984, WO95/00655, WO95/27071, WO95/29993, WO95/34671, WO96/05320, WO94/08026, WO94/11506, WO93/06223, WO94/24299, WO95/14102, WO95/24297, WO95/02697, WO94/28152, WO94/24299, WO95/09241, WO95/25807, WO95/05835, WO94/18922 and WO95/09654. Alternatively, administration of DNA linked to killed adenovirus as described in Curie (1992) Hum. Gene Ther. 3:147-154 may be employed. The gene delivery vehicles of the invention also include adenovirus associated virus (AAV) vectors. Leading and preferred examples of such vectors for use in this invention are the AAV-2 based vectors disclosed in Srivastava, WO93/09239. Most preferred AAV vectors comprise the two AAV inverted terminal repeats in which the native D-sequences are modified by substitution of nucleotides, such that at least 5 native nucleotides

and up to 18 native nucleotides, preferably at least 10 native nucleotides up to 18 native nucleotides, in most preferably 10 native nucleotides are retained and the remaining nucleotides of the D-sequence are deleted or replaced with non-native nucleotides. The native D-sequences of the AAV inverted terminal repeats are sequences of 20 consecutive nucleotides in each AAV inverted terminal repeat (i.e., there is one sequence at each end) which are not involved in HP formation. The non-native replacement nucleotide may be any nucleotide other than the nucleotide found in the native D-sequence in the same position. Other employable exemplary AAV vectors are pWP-19, pWN- 1, both of which are disclosed in Nahreini (1993) Gene 124:257-262. Another example of such an AAV vector is psub201 (see Samulski (1987) J. Virol. 61:3096). Another exemplary AAV vector is the Double-D ITR vector, Construction of the Double-D ITR vector is disclosed in US Patent 5,478, 745. Still other vectors are those disclosed in Carter US Patent 4,797, 368 and Muzyczka US Patent 5,139,941, Chartejee US Patent 5,474,935, and Kotin WO94/288157. Yet a further example of an AAV vector employable in this invention is SSV9AFABTKneo, which contains the AFP enhancer and albumin promoter and directs expression predominantly in the liver. Its structure and construction are disclosed in Su (1996) Human Gene Therapy 7:463-470. Additional AAV gene therapy vectors are described in US 5,354,678, US 5,173,414, US 5,139,941, and US 5,252,479.

**[0231]** The gene therapy vectors of the invention also include herpes vectors. Leading and preferred examples are herpes simplex virus vectors containing a sequence encoding a thymidine kinase polypeptide such as those disclosed in US 5,288,641 and EP0176170 (Roizman). Additional exemplary herpes simplex virus vectors include HFEM/ICP6-LacZ disclosed in WO95/04139 (Wistar Institute), pHSVIac described in Geller (1988) Science 241:1667-1669 and in WO90/09441 and WO92/07945, HSV Us3::pgC-lacZ described in Fink (1992) Human Gene Therapy 3:11-19 and HSV 7134, 2 RH 105 and GAL4 described in EP 0453242 (Breakefield), and those deposited with the ATCC as accession numbers ATCC VR-977 and ATCC VR-260.

**[0232]** Also contemplated are alpha virus gene therapy vectors that can be employed in this invention. Preferred alpha virus vectors are Sindbis viruses vectors. Togaviruses, Semliki Forest virus (ATCC VR-67; ATCC VR- 1247), Middleberg virus (ATCC VR-370), Ross River virus (ATCC VR-373; ATCC VR-1246), Venezuelan equine encephalitis virus (ATCC VR923; ATCC VR- 1250; ATCC VR-1249; ATCC VR-532), and those described in US

patents 5,091, 309, 5,217,879, and WO92/10578. More particularly, those alpha virus vectors described in US Serial No. 08/405,627, filed March 15, 1995, WO94/21792, WO92/10578, WO95/07994, US 5,091,309 and US 5,217,879 are employable. Such alpha viruses may be obtained from depositories or collections such as the ATCC in Rockville, Maryland or isolated from known sources using commonly available techniques. Preferably, alphavirus vectors with reduced cytotoxicity are used (see USSN 08/679640).

**[0233]** DNA vector systems such as eukaryotic layered expression systems are also useful for expressing the nucleic acids of the invention. See WO95/07994 for a detailed description of eukaryotic layered expression systems. Preferably, the eukaryotic layered expression systems of the invention are derived from alphavirus vectors and most preferably from Sindbis viral vectors.

**[0234]** Other viral vectors suitable for use in the present invention include those derived from poliovirus, for example ATCC VR-58 and those described in Evans, Nature 339 (1989) 385 and Sabin (1973) J. Biol. Standardization 1:115; rhinovirus, for example ATCC VR-l 110 and those described in Arnold (1990) J Cell Biochem L401; pox viruses such as canary pox virus or vaccinia virus, for example ATCC VR-I I I and ATCC VR- 2010 and those described in Fisher-Hoch (1989) Proc Natl Acad Sci 86:317; Flexner (1989) Ann NY Acad Sci 569:86, Flexner (1990) Vaccine 8:17; in US 4,603,112 and US 4,769,330 and WO89/01973; SV40 virus, for example ATCC VR-305 and those described in Mulligan (1979) Nature 277:108 and Madzak (1992) J Gen Virol 73:1533; influenza virus, for example ATCC VR-797 and recombinant influenza viruses made employing reverse genetics techniques as described in US 5,166,057 and in Enami (1990) Proc Natl Acad Sci 87:3802-3805; Enami & Palese (1991) J Virol 65:2711-2713 and Luytjes (1989) Cell 59:110, (see also McMichael (1983) NEJ Med 309:13, and Yap (1978) Nature 273:238 and Nature (1979) 277:108); human immunodeficiency virus as described in EP-0386882 and in Buchschacher (1992) J. Virol. 66:2731; measles virus, for example ATCC VR-67 and VR-1247 and those described in EP-0440219; Aura virus, for example ATCC VR-368; Bebaru virus, for example ATCC VR- 600 and ATCC VR-1240; Cabassou virus, for example ATCC VR-922; Chikungunya virus, for example ATCC VR-64 and ATCC VR-1241; Fort Morgan Virus, for example ATCC VR-924; Getah virus, for example ATCC VR-369 and ATCC VR-1243; Kyzylagach virus, for example ATCC VR-927; Mayaro virus, for example ATCC VR-66; Mucambo virus, for example ATCC VR-580 and ATCC VR-1 244; Ndumu virus, for

example ATCC VR-37 1; Pixuna virus, for example ATCC VR-372 and ATCC VR-1245; Tonate virus, for example ATCC VR-925; Triniti virus, for example ATCC VR-469; Una virus, for example ATCC VR-374; Whataroa virus, for example ATCC VR-926; Y-62-33 virus, for example ATCC VR-375; O'Nyong virus, Eastern encephalitis virus, for example ATCC VR-65 and ATCC VR-1242; Western encephalitis virus, for example ATCC VR-70, ATCC VR-1251, ATCC VR-622 and ATCC VR-1252; and coronavirus, for example ATCC VR-740 and those described in Hamre (1966) Proc Soc Exp Biol Med 121:190.

**[0235]** Delivery of the compositions of this invention into cells is not limited to the above mentioned viral vectors. Other delivery methods and media may be employed such as, for example, nucleic acid expression vectors, polycationic condensed DNA linked or unlinked to killed adenovirus alone, for example see US Serial No. 081366,787, filed December 30, 1994 and Curie] (1992) Hum Gene Ther 3:147-154 ligand linked DNA, for example see Wu (1989) J Biol Chem 264:16985-16987, eukaryotic cell delivery vehicles cells, for example see US Serial No. 08/240,030, filed May 9, 1994, and US Serial No. 08/404,796, deposition of photopolymerized hydrogel materials, hand-held gene transfer particle gun, as described in US Patent 5,149,655, ionizing radiation as described in US5,206,152 and in WO92/1 1033, nucleic charge neutralization or fusion with cell membranes. Additional approaches are described in Philip (1994) Mol Cell Biol 14:2411-2418 and in Woffendin (1994) Proc NatlAcad Sci 91:1581-1585.

**[0236]** Particle mediated gene transfer may be employed, for example see US Serial No. 60/023,867. Briefly, the sequence can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, as described in Wu & Wu (1987) J. Biol. Chem. 262:4429- 4432, insulin as described in Hucked (1990) Biochem Pharmacol 40:253-263, galactose as described in Plank (1992) Bioconjugate Chem 3:533-539, lactose or transferrin.

**[0237]** Naked DNA may also be employed. Exemplary naked DNA introduction methods are described in WO 90/11092 and US 5,580,859. Uptake efficiency may be improved using biodegradable latex beads. DNA coated latex beads are efficiently transported into cells after endocytosis initiation by the beads. The method may be improved further by treatment of the

beads to increase hydrophobicity and thereby facilitate disruption of the endosome and release of the DNA into the cytoplasm.

**[0238]** Liposomes that can act as gene delivery vehicles are described in US 5, 422,120, WO95/13796, WO94/23697, WO91/14445 and EP-524,968. As described in US Ser. No. 60/023,867, on non-viral delivery, the nucleic acid sequences encoding a polypeptide can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then be incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, insulin, galactose, lactose, or transferrin, Other delivery systems include the use of liposomes to encapsulate DNA comprising the gene under the control of a variety of tissue-specific or ubiquitously-active promoters. Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin et al (1994) Proc. Natl. Acad. Sci. USA 91(24):11581-11585. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials. Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun, as described in US 5,149,655; use of ionizing radiation for activating transferred gene, as described in US 5,206,152 and WO92/11033. Exemplary liposome and polycationic gene delivery vehicles are those described in US 5,422,120 and 4,762,915; in WO 95/13796; WO94/23697; and WO91/14445; in EP-0524968; and in Stryer, Biochemistry, pages 236-240 (1975) W.H. Freeman, San Francisco; Szoka (1980) Biochem Biophys Acta 600:1; Bayer (1979)Biochem BiophysActa 550:464;Rivnay (1987) Meth Enzymol 149:119;Wang (1987)Proc Natl Acad Sci 84:7851; Plant (1989) Anal Biochem 176:420.

**[0239]** A polynucleotide composition can comprises therapeutically effective amount of a gene therapy vehicle, as the term is defined above. For purposes of the present invention, an effective dose will be from about 0. 01 mg/kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

Delivery Methods

**[0240]** Once formulated, the polynucleotide compositions of the invention can be administered (1) directly to the subject; (2) delivered ex vivo, to cells derived from the

subject; or (3) in vitro for expression of recombinant proteins. The subjects to be treated can be mammals or birds. Also, human subjects can be treated.

**[0241]** Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal or transcutaneous applications (e. g., see WO98/20734), needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

**[0242]** Methods for the ex vivo delivery and reimplantation of transformed cells into a subject are known in the aft and described in e.g., WO93/14778. Examples of cells useful in ex vivo applications include, for example, stem cells, particularly hematopoietic, lymph cells, macrophages, dendritic cells, or tumor cells.

**[0243]** Generally, delivery of nucleic acids for both ex vivo and in vitro applications can be accomplished by the following procedures, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei, all well known in the art.

Polynucleotide and polypeptide pharmaceutical compositions

**[0244]** In addition to the pharmaceutically acceptable carriers and salts described above, the following additional agents can be used with polynucleotide and/or polypeptide compositions.

*i. Polypeptides*

**[0245]** One example are polypeptides which include, without limitation: asioloorosomucoid (ASOR); transferrin; asialoglycoproteins; antibodies; antibody fragments; ferritin; interleukins; interferons, granulocyte, macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), macrophage colony stimulating factor (M-CSF), stem cell factor and erythropoietin. Viral antigens, such as envelope proteins, can also be used. Also, proteins from other invasive organisms, such as the 17 amino acid peptide from the circumsporozoite protein of plasmodium falciparum known as RII

*ii. Hormones, Vitamins, etc.*

**[0246]** Other groups that can be included are, for example: hormones, steroids, androgens, estrogens, thyroid hormone, or vitamins, folic acid.

*iii. Polyalkylenes, Polysaccharides, etc.*

**[0247]** Also, polyalkylene glycol can be included with the desired polynucleotides/polypeptides. In a preferred embodiment, the polyalkylene glycol is polyethlylene glycol. In addition, mono-, di-, or polysaccharides can be included. In a preferred embodiment of this aspect, the polysaccharide is dextran or DEAE-dextran. Also, chitosan and poly(lactide-co-glycolide)

*iv. Lipids, and Liposomes*

**[0248]** The desired polynucleotide/polypeptide can also be encapsulated in lipids or packaged in liposomes prior to delivery to the subject or to cells derived therefrom.

**[0249]** Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed polynucleotide to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight (199 1) Biochim. Biophys. Acta. 1097:1-17; Straubinger (1983) Meth. Enzymol. 101:512-527.

**[0250]** Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Feigner (1987) Proc. Natl. Acad. Sci. USA 84:7413-7416); mRNA (Malone (1989) Proc. Natl. Acad. Sci. USA 86:6077-6081); and purified transcription factors (Debs (1990) J. Biol. Chem. 265:10189-10192), in functional form.

**[0251]** Cationic liposomes are readily available. For example, N(l-2, 3-dioleyloxy)propyl)-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Feigner supra). Other commercially available liposomes include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boerhinger). Other cationic liposomes can be prepared from readily available materials

using techniques well known in the art. See, e.g., Szoka (1978) Proc. Natl. Acad. Sci. USA 75:4194-4198; WO90/11092 for a description of the synthesis of DOTAP (1 2 - bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

**[0252]** Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphoshatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

**[0253]** The liposomes can comprise multilammelar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See e.g., Straubinger (1983) Meth. Immunol. 101:512-527; Szoka (1978) Proc. Nall. Acad. Sci. USA 75:4194-4198; Papahadjopoulos (1975) Biochim. Biophys. Acta 394:483; Wilson (1979) Cell 17:77); Deamer & Bangham (1976) Biochim. Biophys. Acta 443:629; Ostro (1977) Biochem. Biophys. Res. Commun. 76:836; Fraley (1979) Proc. Natl. Acad. Sci. USA 76:3348); Enoch & Strittmatter (1979) Proc. Natl. Acad. Sci. USA 76:145; Fraley (1980) J. Biol. Chem. (1980) 255:10431; Szoka & Papahadjopoulos (1978)Proc. Natl. Acad. Sci. USA 75:145; and Schaefer- Ridder (1982) Science 215:166.

*v. Lipoproteins*

**[0254]** In addition, lipoproteins can be included with the polynucleotide/polypeptide to be delivered. Examples of lipoproteins to be utilized include: chylomicrons, HDL, IDL, LDL, and VLDL. Mutants, fragments, or fusions of these proteins can also be used. Also, modifications of naturally occurring lipoproteins can be used, such as acetylated LDL. These lipoproteins can target the delivery of polynucleotides to cells expressing lipoprotein receptors. Preferably, if lipoproteins are including with the polynucleotide to be delivered, no other targeting ligand is included in the composition.

**[0255]** Naturally occurring lipoproteins comprise a lipid and a protein portion. The protein portion are known as apoproteins. At the present, apoproteins A, B, C, D, and E have been

isolated and identified. At least two of these contain several proteins, designated by Roman numerals, AI, AII, AIV; CI, CII, CIII.

**[0256]** A lipoprotein can comprise more than one apoprotein. For example, naturally occurring chylomicrons comprises of A, B, C, and E, over time these lipoproteins lose A and acquire C and E apoproteins. VLDL comprises A, B, C, and E apoproteins, LDL comprises apoprotein B; and HDL comprises apoproteins A, C, and E. The amino acid of these apoproteins are known and are described in, for example, Breslow (1985) Annu Rev. Biochem 54:699; Law (1986) Adv. Exp Med. Biol. 151:162; Chen (1986) J Biol Chem 261:12918; Kane (1980) Proc Natl Acad Sci USA 77:2465; and Utermann (1984) Hum Genet 65:232.

**[0257]** Lipoproteins contain a variety of lipids including, triglycerides, cholesterol (free and esters), and phospholipids. The composition of the lipids varies in naturally occurring lipoproteins. For example, chylomicrons comprise mainly triglycerides. A more detailed description of the lipid content of naturally occurring lipoproteins can be found, for example, in Meth. Enzymol. 128 (1986). The composition of the lipids are chosen to aid in conformation of the apoprotein for receptor binding activity. The composition of lipids can also be chosen to facilitate hydrophobic interaction and association with the polynucleotide binding molecule.

**[0258]** Naturally occurring lipoproteins can be isolated from serum by ultracentrifugation, for instance. Such methods are described in Meth. Enzymol. (supra); Pitas (1980) J. Biochem. 255:5454-5460 and Mahey (1979) J Clin. Invest 64:743-750. Lipoproteins can also be produced by in vitro or recombinant methods by expression of the apoprotein genes in a desired host cell. See, for example, Atkinson (1986) Annu Rev Biophys Chem 15:403 and Radding (1958)Biochim BiophysActa 30:443. Lipoproteinscan also be purchased from commercial suppliers, such as Biomedical Technologies, Inc., Stoughton, Massachusetts, USA. Further description of lipoproteins can be found in Zuckermann et. al. WO98/06437..

*vi. Polycationic Agents*

**[0259]** Polycationic agents can be included, with or without lipoprotein, in a composition with the desired polynucleotide/polypeptide to be delivered.

**[0260]** Polycationic agents, typically, exhibit a net positive charge at physiological relevant pH and are capable of neutralizing the electrical charge of nucleic acids to facilitate delivery to a desired location. These agents have both *in vitro*, *ex vivo*, and *in vivo* applications. Polycationic agents can be used to deliver nucleic acids to a living subject either intramuscularly, subcutaneously, etc. The following are examples of useful polypeptides as polycationic agents: polylysine, polyarginine, polyornithine, and protamine. Other examples include histones, protamines, human serum albumin, DNA binding proteins, non- histone chromosomal proteins, coat proteins from DNA viruses, such as (X174, transcriptional factors also contain domains that bind DNA and therefore may be useful as nucleic acid condensing agents. Briefly, transcriptional factors such as C/CEBP, c-jun, c-fos, AP-1, AP-2, AP-3, CPF, Prot-1, Sp-1, Oct-1, Oct-2, CREP, and TFIID contain basic domains that bind DNA sequences.

**[0261]** Organic polycationic agents include: spermine, spermidine, and putrescine.

**[0262]** The dimensions and of the physical properties of a polycationic agent can be extrapolated from the list above, to construct other polypeptide polycationic agents or to produce synthetic polycationic agents.

**[0263]** Synthetic polycationic agents which are useful include, for example, DEAE-dextran, polybrene. Lipofectin™, and lipofectAMINE™ are monomers that form polycationic complexes when combined with polynucleotides/polypeptides.

Immunodiagnostic Assays

**[0264]** Another aspect of the present invention includes GBS 80 immunogenic polypeptides of the present invention used in immunoassays to detect antibody levels (or, conversely, anti-*Streptococcal* antibodies can be used to detect antigen levels). Immunoassays based on well defined, recombinant antigens can be developed to replace invasive diagnostics methods. Antibodies to GBS 80 immunogenic polypeptides within biological samples, including for example, blood or serum samples, can be detected. Design of the immunoassays is subject to a great deal of variation, and a variety of these are known in the art. Protocols for the immunoassay may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, fluorescent, chemiluminescent, radioactive, or dye molecules.

Assays which amplify the signals from the probe are also known; examples of which are assays which utilize biotin and avidin, and enzyme labeled and mediated immunoassays, such as ELISA assays.

**[0265]** Kits suitable for immunodiagnosis and containing the appropriate labeled reagents are constructed by packaging the appropriate materials, including the compositions of the invention, in suitable containers, along with the remaining reagents and materials (for example, suitable buffers, salt solutions, etc.) required for the conduct of the assay, as well as suitable set of assay instructions.

Nucleic Acid Hybridisation

**[0266]** "Hybridization" refers to the association of two nucleic acid sequences to one another by hydrogen bonding. Typically, one sequence will be fixed to a solid support and the other will be free in solution. Then, the two sequences will be placed in contact with one another under conditions that favor hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridization; agitation; agents to block the non-specific attachment of the liquid phase sequence to the solid support (Denhardt's reagent or BLOTTO); concentration of the sequences; use of compounds to increase the rate of association of sequences (dextran sulfate or polyethylene glycol); and the stringency of the washing conditions following hybridization. See Sambrook et al. (supra) Volume 2, chapter 9, pages 9.47 to 9.57.

**[0267]** "Stringency" refers to conditions in a hybridization reaction that favor association of very similar sequences over sequences that differ. For example, the combination of temperature and salt concentration should be chosen that is approximately 12° to 20°C below the calculated Tm of the hybrid under study. The temperature and salt conditions can often be determined empirically in preliminary experiments in which samples of genomic DNA immobilized on filters are hybridized to the sequence of interest and then washed under conditions of different stringencies. See Sambrook et al. at page 9.50.

**[0268]** Variables to consider when performing, for example, a Southern blot are (1) the complexity of the DNA being blotted and (2) the homology between the probe and the sequences being detected. The total amount of the fragment(s) to be studied can vary a magnitude of 10, from 0.1 to 1 $\mu$g for a plasmid or phage digest to $10^{-9}$ to $10^{-8}$ g for a single copy gene in a highly complex eukaryotic genome. For lower complexity polynucleotides,

substantially shorter blotting, hybridization, and exposure times, a smaller amount of starting polynucleotides, and lower specific activity of probes can be used. For example, a single-copy yeast gene can be detected with an exposure time of only 1 hour starting with 1 $\mu$g of yeast DNA, blotting for two hours, and hybridizing for 4-8 hours with a probe of $10^8$ cpm/$\mu$g. For a single-copy mammalian gene a conservative approach would start with 10 $\mu$g of DNA, blot overnight, and hybridize overnight in the presence of 10% dextran sulfate using a probe of greater than $10^8$ cpm/$\mu$g, resulting in an exposure time of ~24 hours.

**[0269]** Several factors can affect the melting temperature (Tm) of a DNA- DNA hybrid between the probe and the fragment of interest, and consequently, the appropriate conditions for hybridization and washing. In many cases the probe is not 100% homologous to the fragment. Other commonly encountered variables include the length and total G+C content of the hybridizing sequences and the ionic strength and formamide content of the hybridization buffer. The effects of all of these factors can be approximated by a single equation:

Tm= 81 + 16.6(logCi) + 0.4(%(G + C))-0.6(% formamide) - 600/n - 1.5(% mismatch).

**[0270]** where Ci is the salt concentration (monovalent ions) and n is the length of the hybrid in base pairs (slightly modified from Meinkoth & Wahl (1984) Anal. Biochem. 138: 267-284).

**[0271]** In designing a hybridization experiment, some factors affecting nucleic acid hybridization can be conveniently altered. The temperature of the hybridization and washes and the salt concentration during the washes are the simplest to adjust. As the temperature of the hybridization increases (i.e., stringency), it becomes less likely for hybridization to occur between strands that are nonhomologous, and as a result, background decreases. If the radiolabeled probe is not completely homologous with the immobilized fragment (as is frequently the case in gene family and interspecies hybridization experiments), the hybridization temperature must be reduced, and background will increase. The temperature of the washes affects the intensity of the hybridizing band and the degree of background in a similar manner. The stringency of the washes is also increased with decreasing salt concentrations.

**[0272]** In general, convenient hybridization temperatures in the presence of 50% formamide are 42°C for a probe with is 95% to 100% homologous to the target fragment, 37°C for 90% to 95% homology, and 32°C for 85% to 90% homology. For lower

homologies, formamide content should be lowered and temperature adjusted accordingly, using the equation above. If the homology between the probe and the target fragment are not known, the simplest approach is to start with both hybridization and wash conditions which are nonstringent. If non-specific bands or high background are observed after autoradiography, the filter can be washed at high stringency and re-exposed. If the time required for exposure makes this approach impractical, several hybridization and/or washing stringencies should be tested in parallel.

Combinations including GBS 80

**[0273]** Another aspect of the present invention includes combination of one or more of the immunogenic polypeptides with other GBS antigens. Preferably, the combination of GBS antigens consists of three, four, five, six, seven, eight, nine, or ten GBS antigens. Still more preferably, the combination of GBS antigens consists of three, four, or five GBS antigens. Such combinations may include full length and/or antigenic fragments of the respective antigens and include combinations where the polypeptides and antigens are physically linked to one another and combinations where the polypeptides and antigens are not physically linked but are included in the same composition.

**[0274]** Preferably, the combinations of the invention provide for improved immunogenicity over the immunogenicity of the antigens when administered alone. Improved immunogenicity may be measured, for example, by the Active Maternal Immunization Assay. As discussed in Example 1, this assay may be used to measure serum titers of the female mice during the immunization schedule as well as the survival time of the pups after challenge. Preferably, immunization with the immunogenic compositions of the invention yield an increase of at least 2 percentage points (preferably at least 3, 4 or 5 percentage points) in the percent survival of the challenged pups as compared to the percent survival from maternal immunization with a single antigen of the composition when administered alone. Preferably, the increase is at least 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 percentage points. Preferably, the GBS combinations of the invention comprising GBS 80 demonstrate an increase in the percent survival as compared to the percent survival from immunization with a non-GBS 80 antigen alone.

**[0275]** According to one embodiment of the invention, combinations of antigens or fusion proteins containing a portion or portions of the antigens will include GBS 80 or a portion

thereof in combination with from one to 10 antigens, preferably one to 10 or less antigens. Examples of GBS antigens may be found in U.S. Serial Number 10/415,182, filed April 28, 2003, the International Applications (WO04/041157 and WO05/028618), and WO04/099242, each of which is hereby incorporated in its entirety.

GBS polysaccharides

**[0276]** The compositions of the invention may be further improved by including GBS polysaccharides. Preferably, the GBS antigen and the saccharide each contribute to the immunological response in a recipient. The combination is particularly advantageous where the saccharide and polypeptide provide protection from different GBS serotypes.

**[0277]** The combined antigens may be present as a simple combination where separate saccharide and polypeptide antigens are administered together, or they may be present as a conjugated combination, where the saccharide and polypeptide antigens are covalently linked to each other.

**[0278]** Thus the invention provides an immunogenic composition comprising (i) one or more GBS polypeptide antigens and (ii) one or more GBS saccharide antigens. The polypeptide and the polysaccharide may advantageously be covalently linked to each other to form a conjugate.

**[0279]** Between them, the combined polypeptide and saccharide antigens preferably cover (or provide protection from) two or more GBS serotypes (*e.g.* 2, 3, 4, 5, 6, 7, 8 or more serotypes). The serotypes of the polypeptide and saccharide antigens may or may not overlap. For example, the polypeptide might protect against serogroup II or V, while the saccharide protects against either serogroups Ia, Ib, or III. Preferred combinations protect against the following groups of serotypes: (1) serotypes Ia and Ib, (2) serotypes Ia and II, (3) serotypes Ia and III, (4) serotypes Ia and IV, (5) serotypes Ia and V, (6) serotypes Ia and VI, (7) serotypes Ia and VII, (8) serotypes Ia and VIII, (9) serotypes Ib and II, (10) serotypes Ib and III, (11) serotypes Ib and IV, (12) serotypes Ib and V, (13) serotypes Ib and VI, (14) serotypes Ib and VII, (15) serotypes Ib and VIII, 16) serotypes II and m, (17) serotypes II and IV, (18) serotypes II and V, (19) serotypes II and VI, (20) serotypes II and III, (21) serotypes II and VII, (22) serotypes III and IV, (23) serotypes III and V, (24) serotypes III and VI, (25) serotypes III and VII, (26) serotypes III and VIII, (27) serotypes IV and V, (28) serotypes IV and VI, (29) serotypes IV and VII, (30) serotypes IV and VIII,

(31) serotypes V and VI, (32) serotypes V and VII, (33) serotypes V and VIII, (34) serotypes VI and VII, (35) serotypes VI and VIII, and (36) serotypes VII and VIII.

**[0280]** Still more preferably, the combinations protect against the following groups of serotypes: (1) serotypes Ia and II, (2) serotypes Ia and V, (3) serotypes Ib and II, (4) serotypes Ib and V, (5) serotypes III and II, and (6) serotypes III and V. Most preferably, the combinations protect against serotypes III and V. Protection against serotypes II and V is preferably provided by polypeptide antigens.

**[0281]** Protection against serotypes Ia, Ib and/or III may be polypeptide or saccharide antigens.

In one embodiment, the immunogenic composition comprises a GBS saccharide antigen and at least two GBS polypeptide antigens or fragments thereof, wherein said GBS saccharide antigen comprises a saccharide selected from GBS serotype Ia, Ib, and III, and wherein said GBS polypeptide antigens comprise a combination of at least two polypeptide or a fragment thereof selected from the antigen group consisting of GBS 80 (gi:2253618), GBS 67 (gi22537555), SAN1518 (Spb1, gi:77408651), GBS 104 and GBS 322 (the foregoing antigens are described in U.S. Patent App. No. 11/192,046, which is hereby incorporated by reference for all that it teaches and in particular for the antigens and fragments thereof). Preferably, the combination includes GBS 80 or a fragment thereof.

Further antigens

**[0282]** The compositions of the invention may further comprise one or more additional non-GBS antigens, including additional bacterial, viral or parasitic antigens.

**[0283]** In another embodiment, the GBS antigen combinations of the invention are combined with one or more additional, non-GBS antigens suitable for use in a vaccine designed to protect elderly or immunocomprised individuals. For example, the GBS antigen combinations may be combined with an antigen derived from the group consisting of *Enterococcus faecalis*, *Staphylococcus aureus*, *Staphylococcus epidermis*, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *Listeia monocytogenes*, *Neisseria meningitides*, influenza, and Parainfluenza virus ('PIV').

**[0284]** Where a saccharide or carbohydrate antigen is used, it is preferably conjugated to a carrier protein in order to enhance immunogenicity {*e.g.* refs. 42 to 51}. Preferred carrier proteins are bacterial toxins or toxoids, such as diphtheria or tetanus toxoids. The CRM97 diphtheria toxoid is particularly preferred {52}. Other carrier polypeptides include the *N.meningitidis* outer membrane protein {53}, synthetic peptides {54, 55}, heat shock proteins {56, 57}, pertussis proteins {58, 59}, protein D from *H.influenzae* {60}, cytokines {61}, lymphokines, hormones, growth factors, toxin A or B from *C.difficile* {62}, iron uptake proteins {63}, *etc.* Where a mixture comprises capsular saccharides from both serogroups A and C, it may be preferred that the ratio (w/w) of MenA saccharide:MenC saccharine is greater than 1 (*e.g.* 2:1, 3:1, 4:1, 5:1, 10:1 or higher). Different saccharides can be conjugated to the same or different type of carrier protein. Any suitable conjugation reaction can be used, with any suitable linker where necessary.

**[0285]** Toxic protein antigens may be detoxified where necessary *e.g.* detoxification of pertussis toxin by chemical and/or genetic means.

**[0286]** Where a diphtheria antigen is included in the composition it is preferred also to include tetanus antigen and pertussis antigens. Similarly, where a tetanus antigen is included it is preferred also to include diphtheria and pertussis antigens. Similarly, where a pertussis antigen is included it is preferred also to include diphtheria and tetanus antigens.

**[0287]** Antigens in the composition will typically be present at a concentration of at least 1μg/ml each. In general, the concentration of any given antigen will be sufficient to elicit an immune response against that antigen.

**[0288]** As an alternative to using protein antigens in the composition of the invention, nucleic acid encoding the antigen may be used {*e.g.* refs. 64 to 72}. Protein components of the compositions of the invention may thus be replaced by nucleic acid (preferably DNA *e.g.* in the form of a plasmid) that encodes the protein.

## EXAMPLES

Example 1

**[0289]** As described in WO05/028618, both an Active Maternal Immunization Assay and a Passive Maternal Immunization Assay were conducted on fragments of the GBS 80 protein.

**[0290]** As used herein, an Active Maternal Immunization assay refers to an *in vivo* protection assay where female mice are immunized with the test antigen composition. The female mice are then bred and their pups are challenged with a lethal dose of GBS. Serum titers of the female mice during the immunization schedule are measured as well as the survival time of the pups after challenge.

**[0291]** Specifically, the Active Maternal Immunization assays referred to herein used groups of four CD-1 female mice (Charles River Laboratories, Calco Italy). These mice were immunized intraperitoneally with the selected proteins in Freund's adjuvant at days 1, 21 and 35, prior to breeding. 6-8 weeks old mice received 20 μg protein/dose when immunized with a single antigen, 30-45 μg protein/dose (15 μg each antigen) when immunized with combination of antigens. The immune response of the dams was monitored by using serum samples taken on day 0 and 49. The female mice were bred 2-7 days after the last immunization (at approximately t=36-37), and typically had a gestation period of 21 days. Within 48 hours of birth, the pups were challenged via I.P. with GBS in a dose approximately equal to an amount which would be sufficient to kill 70-90 % of unimmunized pups (as determined by empirical data gathered from PBS control groups). The GBS challenge dose is preferably administered in 50μl of THB medium. Preferably, the pup challenge takes place at 56 to 61 days after the first immunization. The challenge inocula were prepared starting from frozen cultures diluted to the appropriate concentration with THB prior to use. Survival of pups was monitored for 5 days after challenge.

**[0292]** As used herein, the Passive Maternal Immunization Assay refers to an *in vivo* protection assay where pregnant mice are passively immunized by injecting rabbit immune sera (or control sera) approximately 2 days before delivery. The pups are then challenged with a lethal dose of GBS.

**[0293]** Specifically, the Passive Maternal Immunization Assay referred to herein used groups of pregnant CD1 mice which were passively immunized by injecting 1 ml of rabbit immune sera or control sera via I.P., 2 days before delivery. Newborn mice (24-48 hrs after birth) are challenged via I.P. with a 70-90% lethal dose of GBS serotype III COH1. The challenge dose, obtained by diluting a frozen mid log phase culture, was administered in 50μl of THB medium.

[0294] For both assays, the number of pups surviving GBS infection was assessed every 12 hrs for 4 days. Statistical significance was estimated by Fisher's exact test.

[0295] The results of each assay for immunization with SEQ ID NO: 5, SEQ ID NO: 6, PBS and GBS whole cell are set forth in Tables 1 and 2 below.

| TABLE 1: Active Maternal Immunization | | | |
|---|---|---|---|
| Antigen | Alive/total | %Survival | Fisher's exact test |
| PBS (neg control) | 13/80 | 16% | |
| GBS (whole cell) | 54/65 | 83% | P<0.00000001 |
| GBS 80 (intact) | 62/70 | 88% | P<0.0000001 |
| GBS 80 (fragment) SEQ ID5 | 35/64 | 55% | P=0.0000013 |
| GBS 80 (fragment) SEQ ID6 | 13/67 | 19% | P=0.66 |

| Table 2: Passive Maternal Immunization | | | |
|---|---|---|---|
| Antigen | Alive/total | %Survival | Fisher's exact test |
| PBS (neg control) | 12/42 | 28% | |
| GBS (whole cell) | 48/52 | 92% | P<0.0000001 |
| GBS 80 (intact) | 48/55 | 87% | P<0.00000001 |
| GBS 80 (fragment) SEQ ID5 | 45/57 | 79% | P=0.0000006 |
| GBS 80 (fragment) SEQ ID6 | 13/54 | 24% | P=1 |

[0296] As shown in Tables 1 and 2, immunization with the SEQ ID NO: 5 GBS 80 fragment provided a substantially improved survival rate for the challenged pups than the comparison SEQ ID NO: 6 GBS 80 fragment. These results indicate that the SEQ ID NO: 5 GBS 80 fragment contains an important immunogenic epitope of GBS 80.

Example 2

[0297] Epitope mapping was conducted to identify the immunogenic polypeptides of the present invention. First, GBS 80 was subject to total digestion with the Asp-N. Figures 1 and 2 show the predicted fragments and their size on MALDI-TOF. Representative conditions for total digestion with Asp-N were:

- Add 0.1% RapiGest SF (Waters) to 100 µl of GBS80 lot F (1.7 µg/µl) and heat at 98°C for 7 minutes.
- Add 2 µg of Endoproteinase Asp-N (Roche) reconstituted in 5 µl double distilled water.

- Incubate at 37°C for 2 hours.
- Add 0.2% formic acid to stop the digestion.
- Store at –20°C.

After total digestion, the peptides were separated by reverse phase chromatography. The identity of each purified peptide was assessed by MALDI-TOF

**[0298]** The specific immunogenic polypeptides were identified by mapping epitopes found within the GBS 80 protein using six different mouse monoclonal antibodies that specifically bind to the GBS 80 protein. Three monoclonal antibodies were identified from a pool of Hybridoma generated by immunizing a mouse with full-length GBS 80: 9A4/77, 19G4/78, and 19F6/77. Three additional antibodies were identified from a pool of Hybridoma generated by immunizing a mouse also with full-length GBS 80: M3/88, M1/77, and M2/77. Figures 3 and 4 summarize the results of FACs analysis and western blots with the six monoclonal antibodies.

**[0299]** The GBS 80 protein was subject to partial digestion with either Asp-N or Arg-C. Representative conditions for partial digestion with Asp-N were as described above. Representative conditions for partial digestion with Arg-C were:

- Added 0.1% RapiGest SF (Waters) to 100 µl of GBS80 lot F (1.7 µg/µl) and to 100 µl of GBS80 lot 3 (2 µg/µl) and heated at 98°C for 7 minutes.
- Added 2.5 µg of Endoproteinase Arg-C (Roche) to each sample reconstituted in 25 µl double distilled water.
- Incubated at 37°C for 2 hours.
- Added 0.2% formic acid to stop the digestion.
- Stored at –20°C.

**[0300]** The partial digests were then run on SDS PAGE. The gels were stained with Coomassie Blue and the individual bands were isolated and subject to MALDI-TOF to

determine the identity of each band on the gel (See e.g., Figure 5 showing an example of identification of the bands produced by partial digestion of GBS 80 with Asp-N).

- Excise bands of interest from the acrylamide gel and transfer to clean Eppendorf tubes.
- Add 100 μl of destaining solution (50% acetonitrile/ 50 mM ammonium bicarbonate) and allow the gel pieces to detain by shaking the tubes.
- Remove destaining solution and wash with 20 μl of acetonitrile.
- Dry the gel pieces.
- Cover the gel pieces with 12 μl of digestion solution (10 μg/ml Promega Trypsin in 50 mM ammonium bicarbonate).
- Incubate at 37°C for 2 hours.
- Transfer the digestion solution to clean Eppendorf tubes and add 5 μl of 0.1% TFA.
- Purify the tryptic peptides with MAP and analyze with MALDI-TOF.

Figure 7 shows an SDS PAGE gel stained with Coomassie blue comparing partial digests of GBS 80 with and without boiling to denature GBS 80 and two different proteases. GBS 80 F and GBS 80 3 represent different conformer of GBS 80 which may be purified from one another and have different protease sensitivities as shown in Figure 7. Figure 8 shows a representative western blot of the SDS PAGE gel shown in Figure 7. As expected, the monoclonal antibody generated with the N-terminal portion of GBS 80 shows a distinct pattern as compared to the monoclonal antibody generated with the C-terminal portion of GBS 80. Figure 9 shows an SDS PAGE gel of the partial digest produced from boiled samples of GBS 80 with the identity of the protein fragments on the right side of the figure as determined with MALDI-TOF.

**[0301]** From the pattern of bands produced on western blot such as on Figure 6 and 8, the epitopes bound by the antibodies were identified. Figure 10 summarizes the results of the western blots. The fifteen of the sixteen fragments from the SDS PAGE gel shown in Figure

9 are displayed as horizontal bars. The pattern of bands observed in western blots is shown along the left with a (+) for each band observed and a (–) for each band missing. The N column corresponds to binding by the monoclonal antibody 9A4/77 and the C column corresponds to binding by the monoclonal antibody M3/88. The circle on the right indicates the epitope for 9A4/77 and the circle on the left indicates the epitope for M3/88. Figure 11 shows the sequence of GBS 80 with the epitope for 9A4/77 highlighted in yellow and the epitope for M3/88 highlighted in light blue with the core in green.

**[0302]** Additional epitope mapping with the other four monoclonal antibodies produced similar results. Representative western blots for the other four monoclonal antibodies are shown in Figure 12. The results are summarized in Figure 13, which shows the sequence of GBS 80 with the epitope for the N-terminal monoclonal antibodies (9A4/77, 19G4/78 and 19F6/77) highlighted in yellow and the epitope for M3/88 highlighted in light blue with the core in green. The C-terminal epitope is the same, while the N-terminal epitope is a bit more extensive than the epitope for 9A4/77 alone. The results of this Example 2 demonstrate that GBS 80 contains at least three immunogenic polypeptides corresponding to amino acids: 54-118, 38-118, and 321-350.

Example 3

**[0303]** Additional epitope mapping was conducted to identify immunogenic polypeptides of the present invention using peptide arrays. A RepliTope™ peptide microarray (JPT Peptide Technologies) was procured which had a series of overlapping peptide fragments of GBS 80 affixed to it in triplicate. The peptide fragments listed in Table 3 were arranged on the microarray slide as shown in Figure 14. The pattern shown in Figure 14 was replicated three times on the microarray slide.

**Table 3: Peptide sequences on the MicroArray**

| Position on the MicroArray | Peptide Sequence | SEQ ID NO: |
|---|---|---|
| 1 | DAAFLEIPVASTI | 13 |
| 2 | FLEIPVASTINEK | 14 |
| 3 | IPVASTINEKAVL | 15 |
| 4 | ASTINEKAVLGKA | 16 |
| 5 | INEKAVLGKAIEN | 17 |
| 6 | KAVLGKAIENTFE | 18 |
| 7 | LGKAIENTFELQY | 19 |
| 8 | AIENTFELQYDHT | 20 |

| | | |
|---|---|---|
| 9 | NTFELQYDHTPDK | 21 |
| 10 | ELQYDHTPDKADN | 22 |
| 11 | YDHTPDKADNPKP | 23 |
| 12 | TPDKADNPKPSNP | 24 |
| 13 | KADNPKPSNPPRK | 25 |
| 14 | NPKPSNPPRKPEV | 26 |
| 15 | PSNPPRKPEVHTG | 27 |
| 16 | PPRKPEVHTGGKR | 28 |
| 17 | KPEVHTGGKRFVK | 29 |
| 18 | VHTGGKRFVKKDS | 30 |
| 19 | GGKRFVKKDSTET | 31 |
| 20 | RFVKKDSTETQTL | 32 |
| 21 | KKDSTETQTLGGA | 33 |
| 22 | STETQTLGGAEFD | 34 |
| 23 | TQTLGGAEFDLLA | 35 |
| 24 | LGGAEFDLLASDG | 36 |
| 25 | AEFDLLASDGTAV | 37 |
| 26 | DLLASDGTAVKWT | 38 |
| 27 | LASDGTAVKWTDA | 39 |
| 37 | MAEVSQERPAKTT | 40 |
| 38 | VSQERPAKTTVNI | 41 |
| 39 | ERPAKTTVNIYKL | 42 |
| 40 | AKTTVNIYKLQAD | 43 |
| 41 | TVNIYKLQADSYK | 44 |
| 42 | IYKLQADSYKSEI | 45 |
| 43 | LQADSYKSEITSN | 46 |
| 44 | DSYKSEITSNGGI | 47 |
| 45 | KSEITSNGGIENK | 48 |
| 46 | ITSNGGIENKDGE | 49 |
| 47 | NGGIENKDGEVIS | 50 |
| 48 | IENKDGEVISNYA | 51 |
| 49 | KDGEVISNYAKLG | 52 |
| 50 | EVISNYAKLGDNV | 53 |
| 51 | SNYAKLGDNVKGL | 54 |
| 52 | AKLGDNVKGLQGV | 55 |
| 53 | GDNVKGLQGVQFK | 56 |
| 54 | VKGLQGVQFKRYK | 57 |
| 55 | LQGVQFKRYKVKT | 58 |
| 56 | VQFKRYKVKTDIS | 59 |
| 57 | KRYKVKTDISVDE | 60 |
| 58 | KVKTDISVDELKK | 61 |
| 59 | TDISVDELKKLTT | 62 |
| 60 | SVDELKKLTTVEA | 63 |
| 61 | ELKKLTTVEAADA | 64 |
| 62 | KLTTVEAADAKVG | 65 |
| 63 | TVEAADAKVGTIL | 66 |
| 64 | AADAKVGTILEEG | 67 |
| 65 | AKVGTILEEGVSL | 68 |

| | | |
|---|---|---|
| 66 | GTILEEGVSLPQK | 69 |
| 67 | LEEGVSLPQKTNA | 70 |
| 68 | GVSLPQKTNAQGL | 71 |
| 69 | LPQKTNAQGLVVD | 72 |
| 70 | KTNAQGLVVDALD | 73 |
| 71 | AQGLVVDALDSKS | 74 |
| 72 | LVVDALDSKSNVR | 75 |
| 73 | DALDSKSNVRYLY | 11 |
| 74 | DSKSNVRYLYVED | 76 |
| 75 | SNVRYLYVEDLKN | 12 |
| 76 | RYLYVEDLKNSPS | 77 |
| 77 | YVEDLKNSPSNIT | 78 |
| 78 | DLKNSPSNITKAY | 79 |
| 79 | NSPSNITKAYAVP | 80 |
| 80 | SPSNITKAYAVPF | 81 |

**[0304]** The peptide microarray slide was used according to the following procedure:

- Treat the slide and the coverslip with a solution of 0.1 mg/ml polyvinylpyrrolidone overnight at 4°C.
- Wash 2 times with $H_2O$ for 5 minutes.
- Incubate with monoclonal antibody 9A4/77 diluted 1:300 in PBS for 1 hour at room temperature.
- Wash 4 times with PBS-Tween 0.1% for 5 minutes.
- Wash 2 times with PBS for 5 minutes.
- Incubate with anti-mouse antibody labeled with Cy5 diluted 1:300 in PBS for 1 hour at room temperature.
- Wash 4 times with PBS-Tween 0.1% for 5 minutes.
- Wash 3 times with $H_2O$ for 5 minutes.
- Dry the slide using a stream of nitrogen.
- Perform fluorescence scans.

**[0305]** The image from the fluorescence scan is shown in figure 15. The control spots are indicated with a dashed circle. The GBS 80 immunogenic polypeptides are indicated with a solid circle. The monoclonal antibody 9A4/77 bound to two polypeptides: SEQ ID NO:11 and SEQ ID NO:13. Figure 16 shows the position of SEQ ID NO:11 relative to the immunogenic polypeptide identified by western blotting (between the blue brackets).

## ADDITIONAL REFERENCES

(the contents of which are hereby incorporated by reference)


1. Tettelin *et al.* (2002) *Proc. Natl. Acad. Sci. USA,* 10.1 073/pnas. 182380799.
2. International patent application WO02/34771.
3. Terpe *et al.*, "Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems", Appl Microbiol Biotechnol (2003) 60:523-533.
4. WO99/27961.
5. WO02/074244.
6. WO02/064162.
7. WO03/028760.
8. Gennaro (2000) *Remington: The Science and Practice of Pharmacy.* 20th ed., ISBN: 0683306472.
9. *Vaccine design: the subunit and adjuvant approach* (1995) Powell & Newman. ISBN 0-306-44867-X.
10. WO00/23105.
11. WO00/07621.
12. Barr, *et al.*, "ISCOMs and other saponin based adjuvants", Advanced Drug Delivery Reviews (1998) *32*:247-271. See also Sjolander, *et al.*, "Uptake and adjuvant activity of orally delivered saponin and ISCOM vaccines", Advanced Drug Delivery Reviews (1998) *32*:321-338.
13. Niikura *et al.*, "Chimeric Recombinant Hepatitis E Virus-Like Particles as an Oral Vaccine Vehicle Presenting Foreign Epitopes", Virology (2002) 293:273-280.
14. Lenz *et al.*, "Papillomarivurs-Like Particles Induce Acute Activation of Dendritic Cells", Journal of Immunology (2001) 5246-5355.
15. Pinto, *et al.*, "Cellular Immune Responses to Human Papillomavirus (HPV)-16 L1 Healthy Volunteers Immunized with Recombinant HPV-16 L1 Virus Like Particles", Journal of Infectious Diseases(2003) *188*:327-338.
16. Gerber *et al.*, "Human Papillomavirus Virus-Like Particles Are Efficient Oral Immunogens when Coadministered with Escherichia coli Heat-Labile Enterotoxin Mutant R192G or CpG", Journal of Virology (2001) *75*(10):4752-4760.
17. Gluck *et al.*, "New Technology Platforms in the Development of Vaccines for the Future", Vaccine (2002) *20*:B10-B16.

18. Johnson *et al.* (1999) *Bioorg Med Chem Lett 9*:2273-2278

19. Meraldi *et al.*, "OM-174, a New Adjuvant with a Potential for Human Use, Induces a Protective Response with Administered with the Synthetic C-Terminal Fragment 242-310 from the circumsporozoite protein of Plasmodium berghei", Vaccine (2003) *21*:2485-2491.

20. Pajak, *et al.*, "The Adjuvant OM-174 induces both the migration and maturation of murine dendritic cells in vivo", Vaccine (2003) *21*:836-842.

21. Kandimalla, *et al.*, "Divergent synthetic nucleotide motif recognition pattern: design and development of potent immunomodulatory oligodeoxyribonucleotide agents with distinct cytokine induction profiles", Nucleic Acids Research (2003) *31*(9): 239 3-2400.

22. Krieg, "CpG motifs: the active ingredient in bacterial extracts?", Nature Medicine (2003) 9(7): 831-835.

23. McCluskie, *et al.*, "Parenteral and mucosal prime-boost immunization strategies in mice with hepatitis B surface antigen and CpG DNA", FEMS Immunology and Medical Microbiology (2002) *32*:179 185.

24. Kandimalla, *et al.*, "Toll-like receptor 9: modulation of recognition and cytokine induction by novel synthetic CpG DNAs", Biochemical Society Transactions (2003) *31* (part 3): 654-658.

25. Blackwell, *et al.*, "CpG-A-Induced Monocyte IFN-gamma-Inducible Protein-10 Production is Regulated by Plasmacytoid Dendritic Cell Derived IFN-alpha", J. Immunol. (2003) *170*(8):4061 4068.

26. Krieg, "From A to Z on CpG", TRENDS in Immunology (2002) 23(2): 64-65.

27. Kandimalla, *et al.*, "Secondary structures in CpG oligonucleotides affect immunostimulatory activity", BBRC (2003) *306*:948-953.

28. Kandimalla, *et al.*, "Toll-like receptor 9: modulation of recognition and cytokine induction by novel synthetic GpG DNAs", Biochemical Society Transactions (2003) *31*(part 3):664-658.

29. Bhagat *et al.*, "CpG penta and hexadeoxyribonucleotides as potent immunomodulatory agents" BBRC (2003) 300:853-861.

30. Singh *et al.* (2001) *J. Cont. Rele. 70*:267-276.

31. WO99/27960.

32. WO99/52549.

33. WO01/21207.

34. WO01/21152.

35. Andrianov *et al.*, "Preparation of hydrogel microspheres by coacervation of aqueous polyphophazene solutions", Biomaterials (1998) *19*(1-3):109-115.

36. Payne *et al.*, "Protein Release from Polyphosphazene Matrices", Adv. Drug. Delivery Review (1998) *31*(3):185-196.

37. Stanley, "Imiqulmod and the imidazoquinolones: mechanism of action and therapeutic potential" Clin Exp Dermatol (2002) *27*(7):571-577.

38. Jones, "Resiquimod 3M", Curr Opin Investig Drugs (2003) *4*(2):214-218.

39. WO99/11241.

40. WO98/57659.

41. European patent applications 0835318, 0735898 and 0761231.

42. Ramsay *et al.* (2001) *Lancet 357*(9251): 195-196.

43. Lindberg (1999) *Vaccine* 17 Suppl 2:S28-36.

44. Buttery & Moxon (2000) *J R Coll Physicians Lond* 34:163-168.

45. Ahmad & Chapnick (1999) *Infect Dis Clin North Am* 13: 113 133, vii.

46. Goldblatt (1998) J. Med. Microbiol. 47:563-567.

47. European patent 0 477 508.

48. US Patent No. 5,306,492.

49. International patent application WO98/42721.

50. *Conjugate Vaccines* (eds. Cruse *et al.*) ISBN 3805549326, particularly vol. 10:48-114.

51. Hermanson (1996) *Bioconjugate Techniques* ISBN: 0123423368 or 012342335X.

52. *Research Disclosure*, 453077 (Jan 2002)

53. EP-A-0372501

54. EP-A-0378881

55. EP-A-0427347

56. WO93/17712

57. WO94/03208

58. WO98/58668

59. EP-A-0471177

60. WO00/56360

61. WO91/01146

62. WO00/61761

63. WO01/72337

64. Robinson & Torres (1997) *Seminars in Immunology* 9:271-283.

65. Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648.

66. Scott-Taylor & Dalgleish (2000) *Expert Opin Investig Drugs* 9:471-480.

67. Apostolopoulos & Plebanski (2000) *Curr Opin Mol. Ther* 2:441-447.

68. Ilan (1999) *Curr Opin Mol. Ther* 1:116-120.

69. Dubensky *et al.* (2000) *Mol. Med* 6:723-732.

70. Robinson & Pertmer (2000) *Adv Virus Res* 55: 1-74.

71. Donnelly *et al.* (2000) *Am J Respir Crit Care Med* 162(4 Pt 2):S190-193.

72. Davis (1999) *Mt. Sinai J. Med.* 66:84-90.

73. *Current Protocols in Molecular Biology* (F.M. Ausubel *et al.*, eds., 1987) Supplement 30.

74. Smith & Waterman (1981) *Rev. Appl. Math.* 2: 482-489.

75. U.S. Patent No. 6,372,223.

76. WO00/15251.

77. WO01/22992.

78. Hehme *et al.* (2004) *Virus Res.* 103(1-2):163-71.

79. U.S. Patent No. 6,355,271.

80. WO00/23105.

81. *Vaccine Design: The Subunit and Adjuvant Approach* (eds. Powell & Newman) Plenum Press 1995 (ISBN 0-306-44867-X).

82. U.S. Patent No. 5,057,540.

83. WO05/02620.

84. WO96/33739.

85. EP-A-0109942.

86. U.S. Pat. No. 4,578,269.

87. WO96/11711.

88. U.S. Pat. No. 6,352,697.

89. WO00/07621 and U.S. Pat. No. 6,506,386.

90. WO04/04762.

91. Barr *et al.* (1998) *Advanced Drug Delivery Reviews* 32:247-271.

92. Sjolanderet *et al* (1998) *Advanced Drug Delivery Reviews* 32:321-338.

93. Pizza *et al.* (2000) *Int J Med Microbiol* 290:455-461.

94. WO95/17211.

95. WO98/42375.

96. Singh *et al.* (2001) *J Cont Release* 70:267-276.

97. WO99/27960.

98. U.S. Pat. No. 6,090,406.

99. U.S. Pat. No. 5,916,588.

100. EP-A-0626169.

101. WO99/52549.

102. WO01/21207.

103. WO01/21152.

104. WO02/72012.

105. Signorelli and Hadden (2003) *Int Immunopharmacol* 3(8):1177-86.

106. WO04/64715.

107. Cooper (1995) *Pharm Biotechnol* 6:559-80.

108. *Vaccine Adjuvants: Preparation Methods and Research Protocols* (Volume 42 of *Methods in Molecular Methods* series). ISBN: 1-59259-083-7. Ed. O'Hagan.

109. WO05/89837.

110. U.S. Pat. No. 6,692,468.

111. WO00/07647.

112. WO99/17820.

113. U.S. Pat. No. 5,971,953.

114. U.S. Pat. No. 4,060,082.

115. EP-A-0520618.

116. WO98/01174.

117. WO90/14837.

118. Podda and Del Giudice (2003) *Expert Rev Vaccines* 2:197-203.

119. Podda (2001) *Vacccine* 19:2673-2680.

120. Allison and Byars (1992) *Res Immunol* 143:519-25.

121. Hariharan *et al.* (1995) *Cancer Res* 55:3486-9.

122. WO95/11700.

123. U.S. Pat. No. 6,080,725.

124. WO05/097181.

125. Han *et al.* (2005) *Impact of Vitamin E on Immune Function and Infectious Diseases in the Aged at Nutrition, Immune functions and Health* EuroConference, Paris, 9-10 June 2005.

126. U.S. Patent No. 6,630,161.

127. WO02/097072.

128. Hayden *et al.* (1998) *J Clin Invest* 101(3):643-9.

129. Tassignon *et al.* (2005) *J Immunol Meth* 305:188-98.

130. Myers *et al.* (1990) pages 145-156 of Cellular and molecular aspects of endotoxin reactions.

131. Ulrich (2000) Chapter 16 (pages 273-282) of reference 108.

132. Johnson *et al.* (1999) *J Med Chem* 42:4640-9.

133. Baldrick *et al.* (2002) *Regulatory Toxicol Pharmacol* 35:398-413.

134. U.S. Pat. No. 4,680,338.

135. U.S. Pat. No. 4,988,815.

136. WO92/15582.

137. Stanley (2002) *Clin Exp Dermatol* 27:571-577.

138. Wu *et al.* (2004) *Antiviral Res.* 64(2):79-83.

139. Vasilakos *et al.* (2000) *Cell Immunol.* 204(1):64-74.

140. U.S. Pat. No. 4689338, 4929624, 5238944, 5266575, 5268376, 5346905, 5352784, 5389640, 5395937, 5482936, 5494916, 5525612, 6083505, 6440992, 6627640, 6656938, 6660735, 6660747, 6664260, 6664264, 6664265, 6667312, 6670372, 6677347, 6677348, 6677349, 6683088, 6703402, 6743920, 6800624, 6809203, 6888000 and 6924293.

141. Jones (2003) *Curr Opin Investig Drugs* 4:214-218.

142. WO2004/060308.

143. WO2004/064759.

144. U.S. Pat. No. 6,924,271.

145. U.S. Patent App. No. 2005/0070556.

146. U.S. Pat. No. 5,658,731.

147. U.S. Pat. No. 5,011,828.

148. WO2004/87153.

149. U.S. Pat. No. 6,605,617.

150. WO02/18383.

151. WO2004/018455.

152. WO03/082272.

153. PCT/US2005/022769.

154. Johnson *et al.* (1999) *Bioorg Med Chem Lett* 9:2273-2278.

155. Evans *et al.* (2003) *Expert Rev Vaccines* 2:219-229.

156. Andrianov *et al.* (1998) *Biomaterials* 19:109-115.

157. Payne *et al.* (1998) *Adv Drug Delivery Review* 31:185-196.

158. Thompson *et al.* (2003) *Methods in Molecular Medicine* 94:255-266.

159. Kandimalla *et al.* (2003) *Nucleic Acids Research* 31:2393-2400.

160. WO02/26757.

161. WO99/62923.

162. Krieg (2003) *Nature Medicine* 9:831-835.

163. McCluskie *et al.* (2002) *FEMS Immunology and Medical Microbiology* 32:179-185.

164. WO98/40100.

165. U.S. Pat. No. 6,207,646.

166. U.S. Pat. No. 6,239,116.

167. U.S. Pat. No. 6,429,199.

168. Kandimalla *et al.* (2003) *Biochemical Society Transactions* 31 (part 3):654-658.

169. Blackwell *et al.* (2003) *J Immunol* 170:4061-4068.

170. Krieg (2002) *Trends Immunol* 23:64-65.

171. WO01/95935.

172. Kandimalla *et al.* (2003) *BBRC* 306:948-953.

173. Bhagat *et al.* (2003) *BBRC* 300:853-861.

174. WO03/035836.

175. WO01/22972.

176. Thompson *et al.* (2005) *J Leukoc Biol* 78: 'The low-toxicity versions of LPS, MPL® adjuvant and RC529, are efficient adjuvants for CD4+ T cells'.

177. UK patent application GB A 2220211.

178. WO 94/21292.

179. WO94/00153.

180. WO95/17210.

181. WO96/26741.

182. WO93/19780.

183. WO03/011223.

184. Meraldi *et al.* (2003) *Vaccine* 21:2485-2491.

185. Pajak *et al.* (2003) *Vaccine* 21:836-842.

186. U.S. Pat. No. 6,586,409.

187. Wong *et al.* (2003) *J Clin Pharmacol* 43(7):735-42.

188. U.S. Pat. App. No. 2005/0215517.

# CLAIMS

WHAT IS CLAIMED IS:

1. A composition comprising a polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:7 and wherein the polypeptide is capable of generating an immune response in a subject.

2. The composition of claim 1 wherein the region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 260 to 30 amino acids.

3. A composition comprising a polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:8 and wherein the polypeptide is capable of generating an immune response in a subject.

4. The composition of claim 3 wherein the region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids.

5. A composition comprising a polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 200 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:9 and wherein the polypeptide is capable of generating an immune response in an animal.

6. The composition of claim 5 wherein the region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 300 to 30 amino acids.

7. A composition comprising a polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 70 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:10 and wherein the polypeptide is capable of generating an immune response in an animal.

8. The composition of claim 5 wherein the region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 70 to 30 amino acids.

9. A composition comprising a polypeptide having a region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2, wherein the region includes SEQ ID

NO:9 and extends no more than 47 amino acids upstream of SEQ ID NO:9, wherein the polypeptide is capable of generating an immune response in an animal.

10. A composition comprising a polypeptide having a region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2, wherein the region includes SEQ ID NO:10 and extends no more than 56 amino acids upstream of SEQ ID NO:10, wherein the polypeptide is capable of generating an immune response in an animal.

11. A composition comprising a polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:11 and wherein the polypeptide is capable of generating an immune response in a subject.

12. The composition of claim 11 wherein the region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids.

13. A composition comprising a polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:12 and wherein the polypeptide is capable of generating an immune response in a subject.

14. The composition of claim 13 wherein the region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids.

15. A vaccine comprising an immunologically effective amount of the composition of one of claims 1-14.

16. The vaccine of claim 15 further comprising a pharmaceutically acceptable carrier.

17. The vaccine of claim 16 wherein the pharmaceutically acceptable carrier is an adjuvant.

18. A method for the therapeutic or prophylactic treatment of GBS infection in an animal susceptible to GBS infection comprising administering to said animal the vaccine of claim

19. A method for the manufacture of a medicament for raising an immune response against GBS comprising combining a composition of one of claims 1-14 with a pharmaceutically effective carrier.

20. A method for detecting a *Streptococcal* infection in an animal comprising:

(a) contacting an antibody containing sample from the animal with a composition of one of claims 1-14; and

(b) detecting an antibody from the antibody containing sample binding to the composition.

21. The method of claim 20 wherein the *Streptococcal* infection is a GBS infection.

22. A nucleic acid encoding a composition of one of claims 1-14.

23. A vector comprising the nucleic acid of claim 22.

24. The vector of claim 23 further comprising a promoter operably linked to the nucleic acid.

25. Use of the compositions of any one of claims 1-14 in the preparation of a medicament for treatment of GBS infection.

26. A method of screening a polypeptide for immunogenicity comprising:

a) providing a polypeptide selected from the group comprising a first polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:7, a second polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:8, a third polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 200 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:9, a fourth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 70 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:10, a fifth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2, wherein the region includes SEQ ID NO:9 and extends no more than 47 amino acids upstream of SEQ ID NO:9, a sixth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2, wherein the region includes SEQ ID NO:10 and extends no more than 56 amino acids

upstream of SEQ ID NO:10, a seventh polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:11 and an eighth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:12;

b) screening the polypeptide for an immunogenic response.

27. The method of claim 26 wherein the first polypeptide's region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 260 to 30 amino acids, the second polypeptide's region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids, the third polypeptide's region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 300 to 30 amino acids, the fourth polypeptide's region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 70 to 30 amino acids, the seventh polypeptide's region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids, and the eighth polypeptide's region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids.

28. The method of claim 26 wherein the screening comprises introducing the polypeptide into an animal, obtaining a sample of immune sera from the animal, and testing the polypeptide for an immunogenic response with the sample of immune sera.

29. An immunogenic polypeptide produced by the process comprising:

a) providing a polypeptide selected from the group comprising a first polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:7, a second polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:8, a third polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 200 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:9, a fourth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 70 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:10, a

fifth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2, wherein the region includes SEQ ID NO:9 and extends no more than 47 amino acids upstream of SEQ ID NO:9, a sixth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2, wherein the region includes SEQ ID NO:10 and extends no more than 56 amino acids upstream of SEQ ID NO:10, a seventh polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:11 and an eighth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:12;

b) screening the polypeptide for an immunogenic response; and

c) identifying the polypeptide as an immunogenic polypeptide based upon the screening.

30. The immunogenic polypeptide of claim 29 wherein the first polypeptide's region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 260 to 30 amino acids, the second polypeptide's region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids, the third polypeptide's region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 300 to 30 amino acids, the fourth polypeptide's region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 70 to 30 amino acids, the seventh polypeptide's region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids, and the eighth polypeptide's region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids.

31. An immunogenic polypeptide produced by the process comprising:

a) providing two or more polypeptides selected from the group comprising a first polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:7, a second polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:8, a third polypeptide having a region of

limited, contiguous sequence identity of at least 80 percent over 200 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:9, a fourth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 70 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:10, a fifth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2, wherein the region includes SEQ ID NO:9 and extends no more than 47 amino acids upstream of SEQ ID NO:9, a sixth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2, wherein the region includes SEQ ID NO:10 and extends no more than 56 amino acids upstream of SEQ ID NO:10, a seventh polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:11 and an eighth polypeptide having a region of limited, contiguous sequence identity of at least 80 percent over 260 or fewer amino acids to SEQ ID NO:2, wherein the region includes SEQ ID NO:12;

b) screening each of the two or more polypeptides for an immunogenic response; and

c) selecting an immunogenic polypeptide from the two or more polypeptides by applying a criterion.

32. The immunogenic polypeptide of claim 31 wherein the first polypeptide's region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 260 to 30 amino acids, the second polypeptide's region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids, the third polypeptide's region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 300 to 30 amino acids, the fourth polypeptide's region of limited, contiguous sequence identity of at least 80 percent to SEQ ID NO:2 is 70 to 30 amino acids, the seventh polypeptide's region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids, and the eighth polypeptide's region of limited, contiguous sequence identity of at least 80 percent SEQ ID NO:2 is 260 to 30 amino acids.

# Figure 1 – Asp-N MS-Digest Predicted Fragments

## Parameters

Database: **User Protein: recombinant GBS80**
Digest Used: **Asp -N**
Max. # Missed Cleavages: **0**
Peptide N terminus: **Hydrogen**
Peptide C terminus: **Free Acid**
User AA Formula 1: **C2 H3 N1 O1**
Cysteine Modification: **acry lamide**
Instrument Name: **MALDI -TOF**
Minimum Digest Fragment Mass: **100**
Maximum Digest Fragment Mass: **60000**
Minimum Digest Fragment Length: **1**
pI of Protein: **5.3**
Protein MW: **53004**
Amino Acid Composition: **A40 D38 E36 F15 G31 H4 I28 K57 L34 M2 N29 P23 Q13 R7 S25 T48 V34 W2 Y18**

| 1 | 11 | 21 | 31 | 41 | 51 | 61 | 71 |
|---|---|---|---|---|---|---|---|
| MAEVSQERPA | KTTVNIYKLQ | ADSYKSEITS | NGGIEN KDGE | VISNYAKL GD | NVKGLQGVQF | KRYKVK TDIS | VDELKKLTTV |
| 81 | 91 | 101 | 111 | 121 | 131 | 141 | 151 |
| EA ADAKVGTI | LEEGVSLPQK | TNAQGLV VDA | LDSKSNVRYL | YV EDLKNSPS | NITKAYAVPF | VLELPVANST | GTGFLSEINI |
| 161 | 171 | 181 | 191 | 201 | 211 | 221 | 231 |
| YPKNVV TDEP | KTDKDVKKLG | QD DAGYTIGE | EFKWFLKSTI | PANL GDYEKF | EI TDKF ADGL | TYKSVGKIKI | GSKTLN RDEH |
| 241 | 251 | 261 | 271 | 281 | 291 | 301 | 311 |
| YT IDEPT VDN | QNTLKITFKP | EKFKEIAELL | KGMTLVKN QD | ALDKATAN TD | DAAFLEIPVA | STINEKAVLG | KAIENTFELQ |
| 321 | 331 | 341 | 351 | 361 | 371 | 381 | 391 |
| YDHT PDK ADN | PKPSNPPRKP | EVHTGGKRFV | KKDSTETQTL | GGAE FDLLA S | DGTAVKW TDA | LIKANTNKNY | IAGEAVTGQP |
| 401 | 411 | 421 | 431 | 441 | 451 | 461 | 471 |
| IKLKSH TDGT | FEIKGLAYA V | DANAEGTAVT | YKLKETKAPE | GYVI PDKEIE | FTVSQTSYNT | KP TDIT VDS A | DAT PDTIKNN |
| 481 | | | | | | | |
| KRP S | | | | | | | |

(SEQ ID NO:4)

# Figure 2 – Predicted Fragments

| Number | m/z (mi) | m/z (av) | Start | End | Missed Cleavages | Database Sequence |
|---|---|---|---|---|---|---|
| 1 | 262.1403 | 262.2876 | 173 | 174 | 0 | (T)DK (D) |
| 1 | 292.1145 | 292.2704 | 468 | 470 | 0 | (V)DSA (D) |
| 1 | 318.1665 | 318.3522 | 280 | 282 | 0 | (Q)DAL (D) |
| 1 | 318.1665 | 318.3522 | 109 | 111 | 0 | (V)DAL (D) |
| 1 | 333.1774 | 333.3669 | 326 | 328 | 0 | (P)DKA (D) |
| 1 | 403.1829 | 403.4149 | 471 | 474 | 0 | (A)DATP (D) |
| 1 | 433.2298 | 433.4850 | 68 | 71 | 0 | (T)DISV (D) |
| 1 | 447.2455 | 447.5121 | 464 | 467 | 0 | (T)DITV (D) |
| 1 | 469.2047 | 469.4777 | 322 | 325 | 0 | (Y)DHTP (D) |
| 1 | 480.2458 | 480.5448 | 214 | 217 | 0 | (T)DKFA (D) |
| 1 | 518.2826 | 518.5914 | 366 | 370 | 0 | (F)DLLAS (D) |
| 1 | 560.2568 | 560.5853 | 244 | 248 | 0 | (I)DEPTV (D) |
| 1 | 589.2833 | 589.6270 | 168 | 172 | 0 | (T)DEPKT (D) |
| 1 | 720.3528 | 720.7622 | 283 | 289 | 0 | (L)DKATANT (D) |
| 1 | 777.3419 | 777.8141 | 238 | 243 | 0 | (R)DEHYTI (D) |
| 1 | 787.4678 | 787.9405 | 175 | 181 | 0 | (K)DVKKLGQ (D) |
| 1 | 877.4420 | 877.9788 | 371 | 378 | 0 | (S)DGTAVKWT (D) |
| 1 | 1044.4890 | 1045.1413 | 206 | 213 | 0 | (G)DYEKFEIT (D) |
| 1 | 1172.6388 | 1173.3227 | 475 | 484 | 0 | (P)DTIKNNKRPS (-) |
| 1 | 1265.6378 | 1266.4023 | 38 | 49 | 0 | (K)DGEVISNYAKLG (D) |
| 1 | 1317.7266 | 1318.5196 | 72 | 83 | 0 | (V)DELKKLTTVEAA (D) |
| 1 | 1355.5967 | 1356.3943 | 353 | 365 | 0 | (K)DSTETQTLGGAEF (D) |
| 1 | 1383.7160 | 1384.5821 | 408 | 420 | 0 | (T)DGTFEIKGLAYAV (D) |
| 1 | 1472.7385 | 1473.6360 | 112 | 123 | 0 | (L)DSKSNVRYLYVE (D) |
| 1 | 1741.8244 | 1742.8486 | 22 | 37 | 0 | (A)DSYKSEITSNGGIENK (D) |
| 1 | 2088.0137 | 2089.2779 | 446 | 463 | 0 | (P)DKEIEFTVSQTSYNTKPT (D) |
| 1 | 2108.1980 | 2109.4953 | 50 | 67 | 0 | (G)DNVKGLQGVQFKRYKVKT (D) |
| 1 | 2178.2610 | 2179.5843 | 218 | 237 | 0 | (A)DGLTYKSVGKIKIGSKTLNR (D) |
| 1 | 2377.2550 | 2378.7705 | 1 | 21 | 0 | (-)MAEVSQERPAKTTVNIYKLQA (D) |
| 1 | 2557.2979 | 2558.9122 | 183 | 205 | 0 | (D)DAGYTIGEEFKWFLKSTIPANLG (D) |
| 1 | 2566.4092 | 2567.9612 | 84 | 108 | 0 | (A)DAKVGTILEEGVSLPQKTNAQGLVV (D) |
| 1 | 2665.3725 | 2667.0071 | 421 | 445 | 0 | (V)DANAEGTAVTYKLKETKAPEGYVIP (D) |
| 1 | 2713.5014 | 2715.1537 | 329 | 352 | 0 | (A)DNPKPSNPPRKPEVHTGGKRFVKK (D) |
| 1 | 3095.6853 | 3097.5644 | 379 | 407 | 0 | (T)DALIKANTNKNYIAGEAVTGQPIKLKSHT (D) |
| 1 | 3394.7786 | 3396.8870 | 291 | 321 | 0 | (D)DAAFLEIPVASTINEKAVLGKAIENTFELQ Y (D) |
| 1 | 3618.9933 | 3621.3046 | 249 | 279 | 0 | (V)DNQNTLKITFKPEKFKEIAELLKGMTLVK NQ (D) |
| 1 | 4721.5082 | 4724.4431 | 124 | 167 | 0 | (E)DLKNSPSNITKAYAVPFVLELPVANSTGT GFLSEINIYPKNVVT (D) |

## Figure 3 – Western Blot and FACS anti GBS80 mAbs results

| | | | | | | Western Blot | | FACS | |
|---|---|---|---|---|---|---|---|---|---|
| α-GBS | Clone | mg | mg/ml | Ig Isotype | | Result | Strain | Result | Strain |
| 80 N-term | **19G4/78** | 4 | 0.9 | G1 | | ++ | JM9130013 | ΔM = 244 | COH1 |
| 80 N-term | **9A4/77** | 5.7 | 1.9 | G2b | | + | JM9130013 | ΔM = 227 | COH1 |
| 80 N-term | **19F6/77** | 2.5 | 0.5 | G1 | | + | JM9130013 | ΔM = 242 | COH1 |
| 80 C-term | **M1/77** | 8.3 | 3.3 | G | | +/- | JM9130013 | ΔM = 182 | COH1 |
| 80 C-term | **M3/88** | 8.5 | 3.4 | G | | +/- | JM9130013 | ΔM = 180 | COH1 |
| 80 C-term | **M2/77** | 8 | 2 | G1 | | +/- | JM9130013 | ΔM = 185 | COH1 |

**Figure 4 - FACS anti GBS80 mAbs**

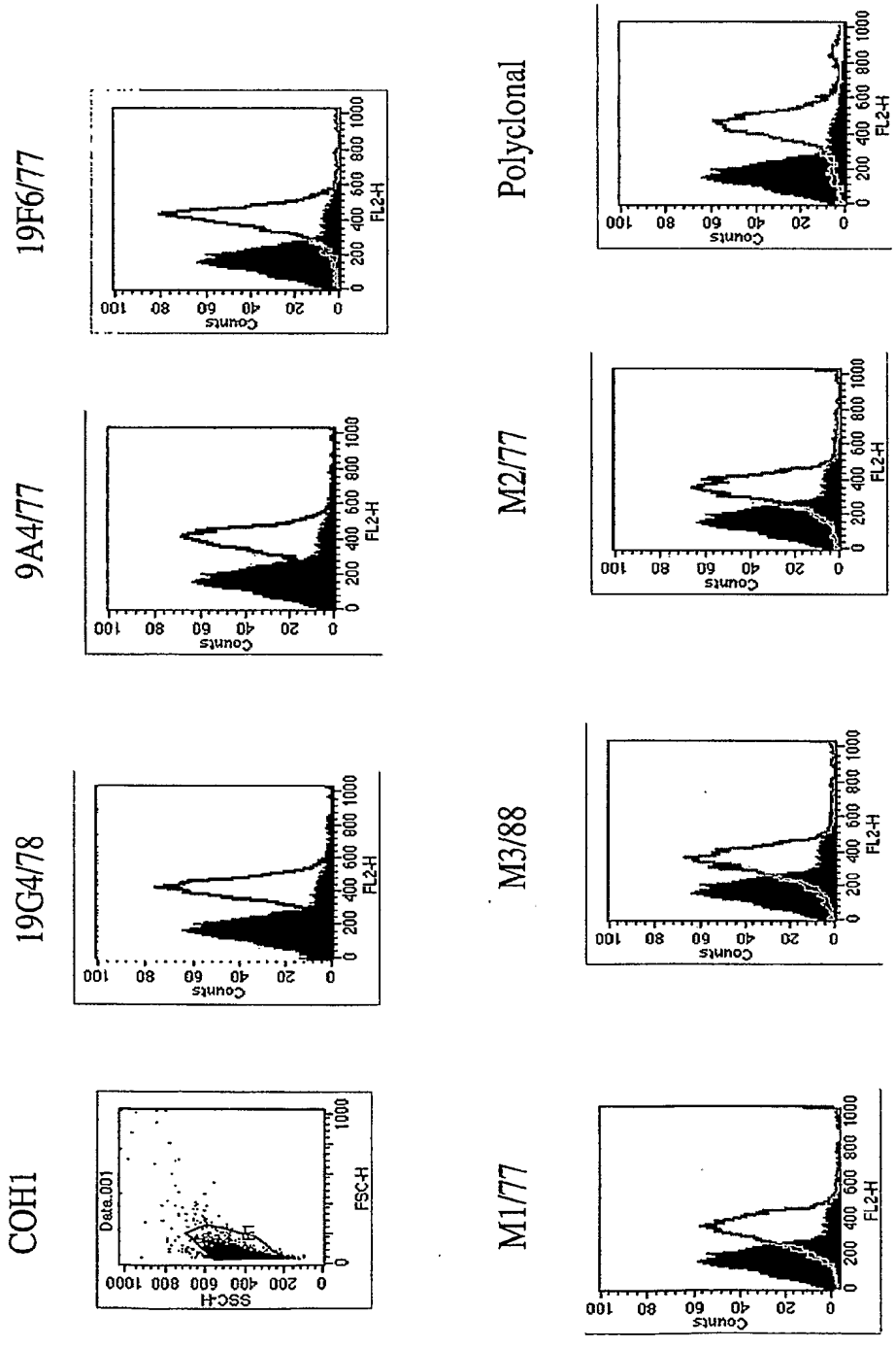


**Figure 5 – MALDI-TOF Identification of Protein Fragments**

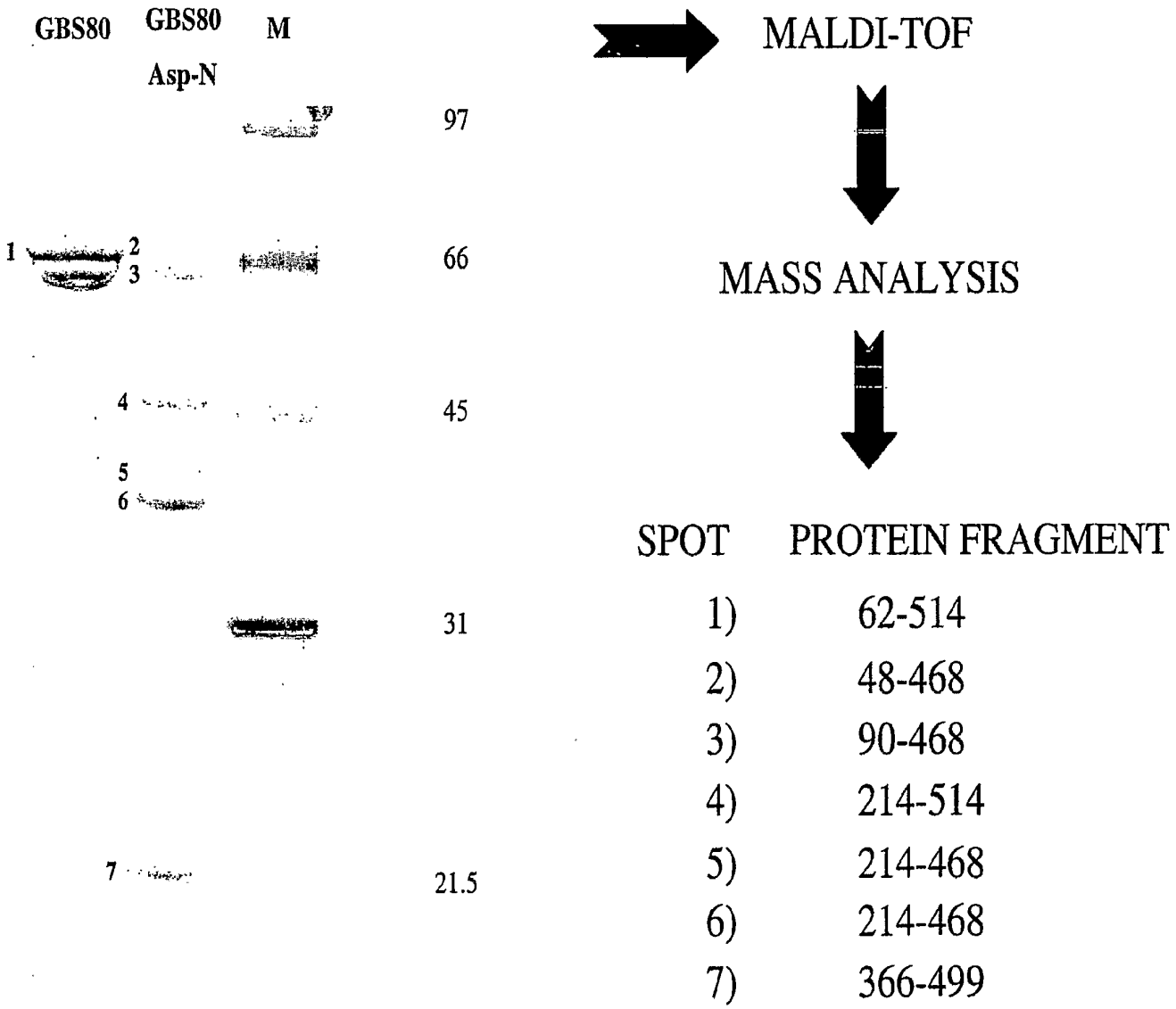


Figure 6 – Representative western blot of Asp-N digest

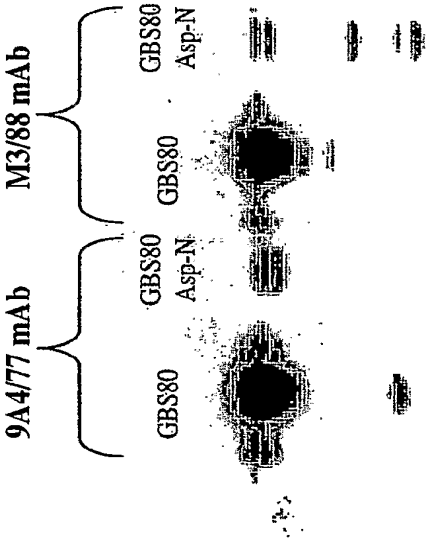
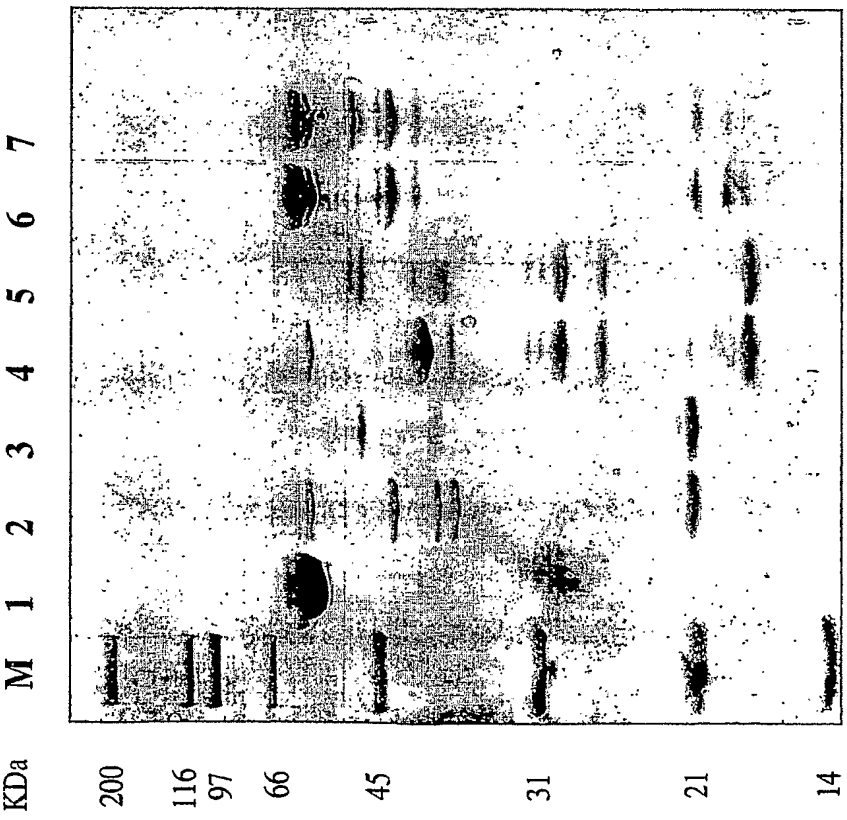
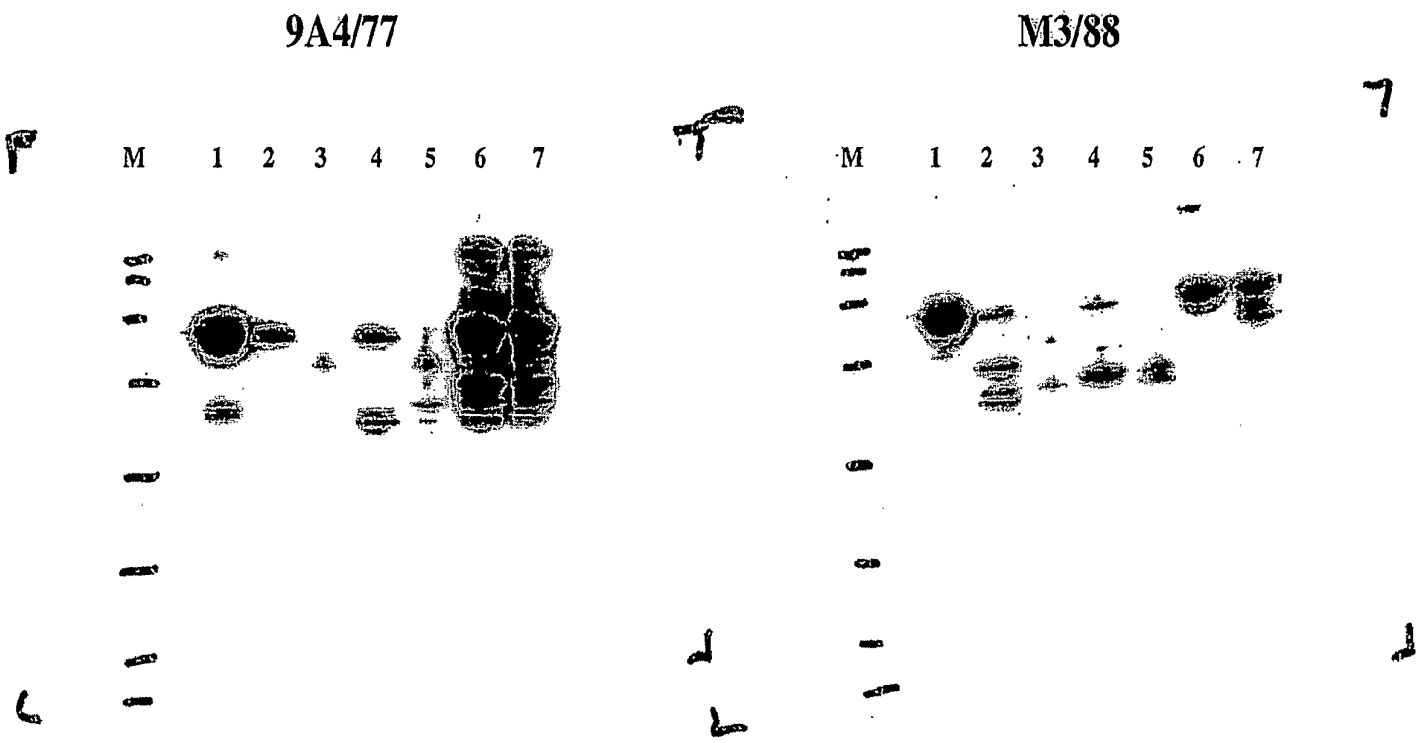


Figure 7 – Representative Coomassie Stain of Digests



## Figure 8 – Representative Western Blot of Digests




**M** Marker

**1** GBS80 F

**2** GBS80 F Asp-N

**3** GBS80 F Asp-N non boiled

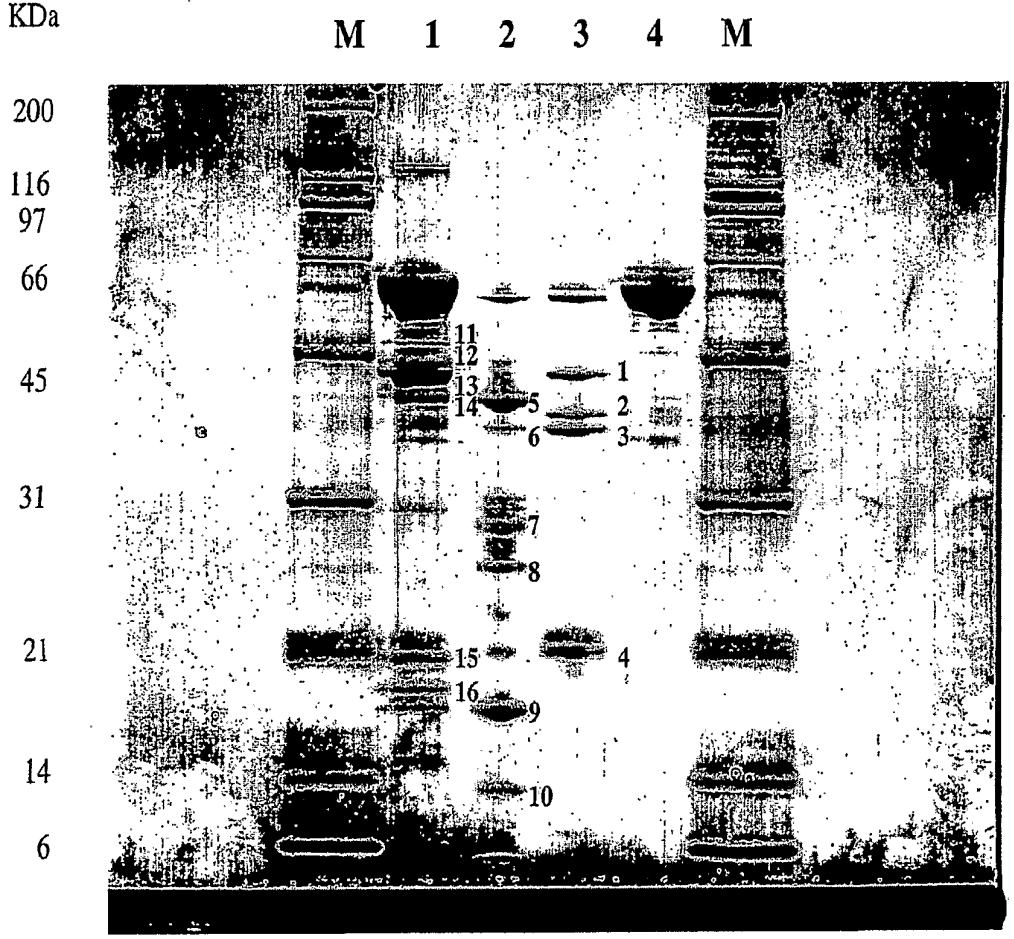
**4** GBS80 F Arg-C

**5** GBS80 F Arg-C non boiled

**6** GBS80 3 Arg-C

**7** GBS80 3 Arg-C non boiled

**Figure 9 – Identification of Protein Fragments from other Digests**




| SPOT | Protein Fragment |
|---|---|
| 1) | 214-514 |
| 2) | 214-468 |
| 3) | 214-468 |
| 4) | 366-499 |
| 5) | 171-514 |
| 6) | 90-342 |
| 7) | 155-374 |
| 8) | 200-374 |
| 9) | 389-514 |
| 10) | not identified |
| 11) | 90-514 |
| 12) | 90-468 |
| 13) | 74-374 |
| 14) | 74-374 |
| 15) | 388-514 |
| 16) | 388-514 |

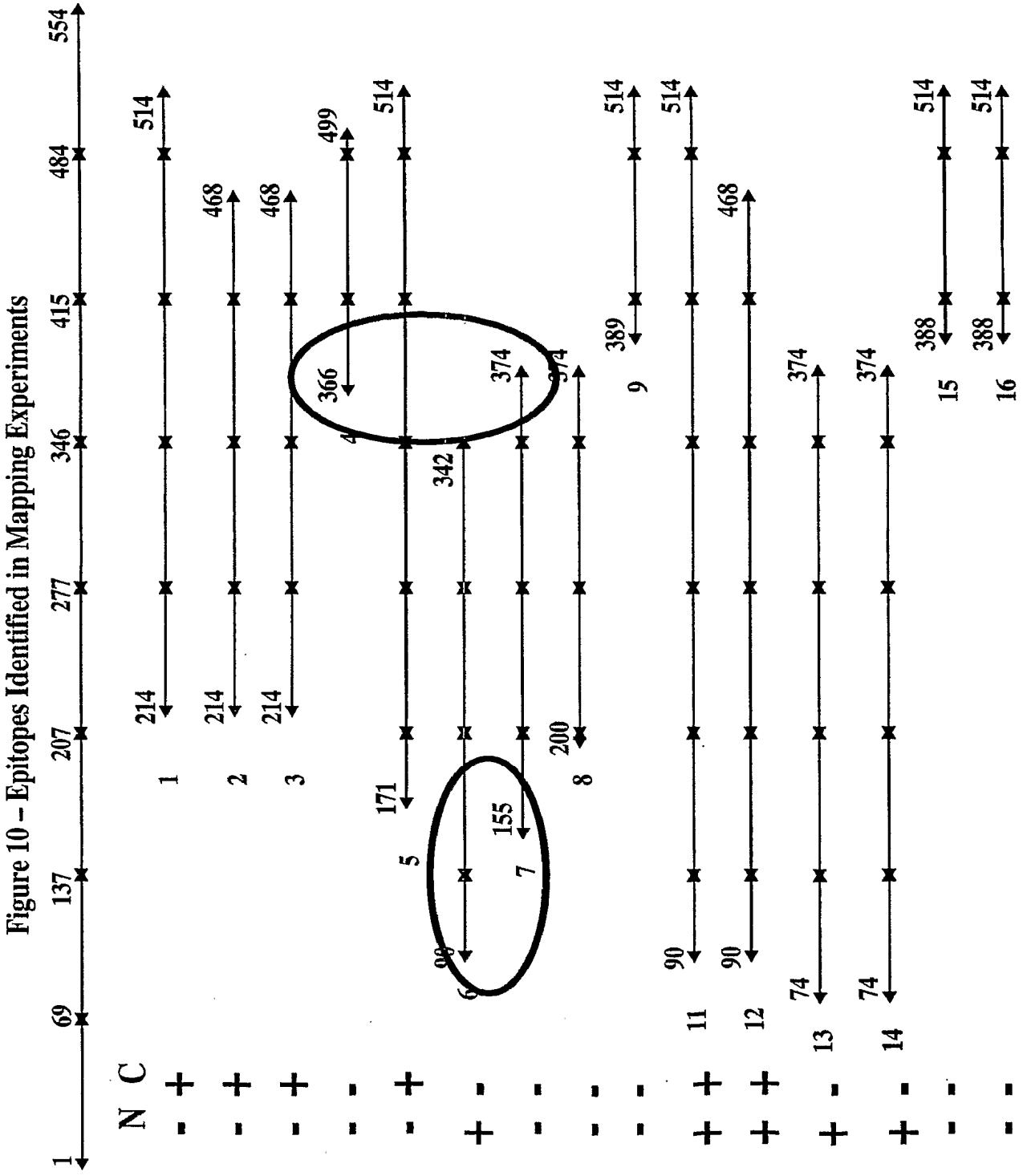


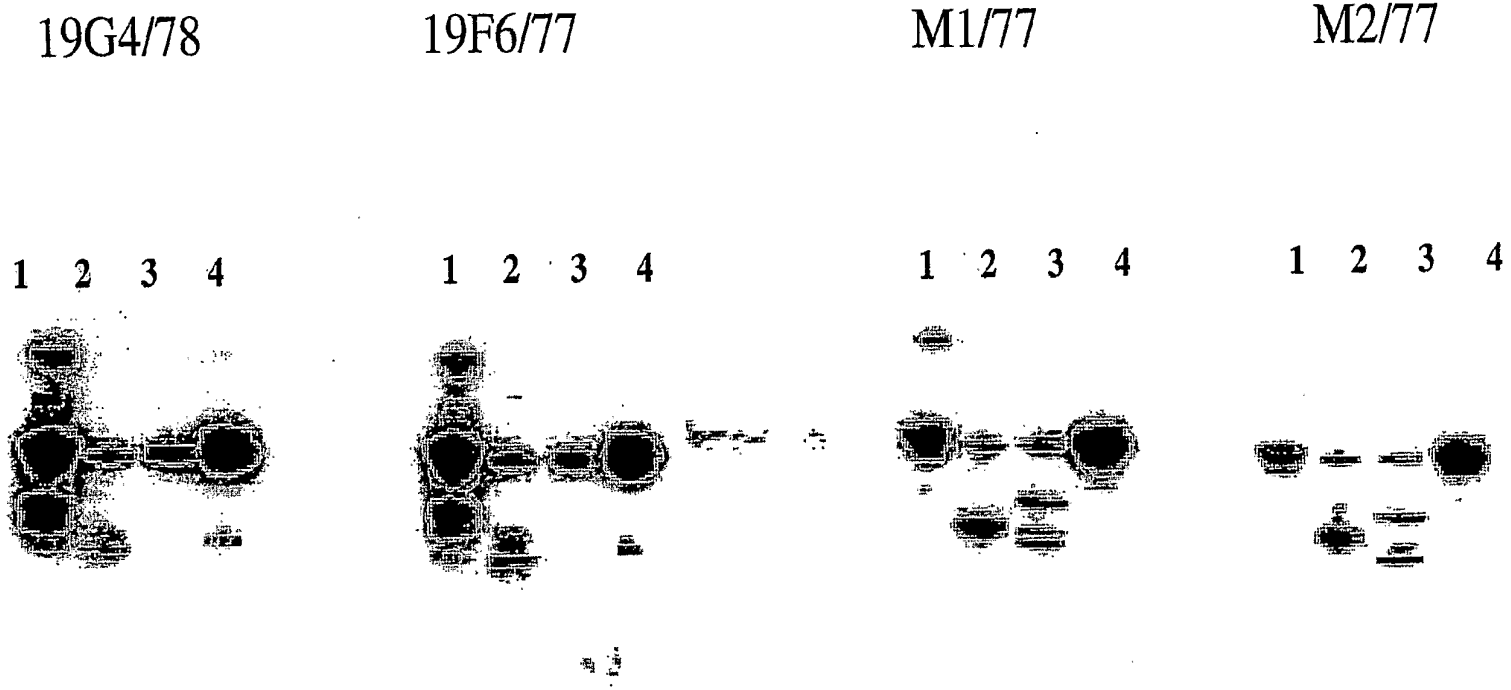

Figure 10 – Epitopes Identified in Mapping Experiments

# Figure 11 – Immunogenic Peptides in GBS 80

| | |
|---|---|
| GBS80 N-term | red |
| GBS80 C-term | blue |
| mAb 9A4/77 on GBS80 | yellow |
| mAb M3/88 on GBS80 | green + light blue |

| 1 | 11 | 21 | 31 | 41 | 51 | 61 | 71 |
|---|---|---|---|---|---|---|---|
| MAEVSQERPA | KTTVNIYKLQ | ADSYKSEITS | NGGIENKDGE | VISNYAKLGD | NVKGLQGVQF | KRYKVKTDIS | VDELKKLTTV |
| 81 | 91 | 101 | 111 | 121 | 131 | 141 | 151 |
| EAADAKVGTI | LEEGVSLPQK | TNAQGLVVDA | LDSKSNVRYL | YVEDLKNSPS | NITKAYAVPF | VLELPVANST | GTGFLSEINI |
| 161 | 171 | 181 | 191 | 201 | 211 | 221 | 231 |
| YPKNVVTDEP | KTDKDVKKLG | QDDAGYTIGE | EFKWFLKSTI | PANLGDYEKF | EITDKFADGL | TYKSVGKIKI | GSKTLNRDEH |
| 241 | 251 | 261 | 271 | 281 | 291 | 301 | 311 |
| YTIDEPTVDN | QNTLKITFKP | EKFKEIAELL | KGMTLVKNQD | ALDKATANTD | DAAFLEIPVA | STINEKAVLG | KAIENTFELQ |
| 321 | 331 | 341 | 351 | 361 | 371 | 381 | 391 |
| VDHIPDKADN | PKPSNEPRKP | EVHIGGKRFV | KKDSTETQTL | GGAEFDLLAS | DGTAVKWTDA | LIKANTNKNY | IAGEAVTGQP |
| 401 | 411 | 421 | 431 | 441 | 451 | 461 | 471 |
| IKLKSHTDGT | FEIKGLAYAV | DANAEGTAVT | YKLKETKAPE | GYVIPDKEIE | FTVSQTSYNT | KPTDITVDSA | DATPDTIKNN |
| 481 | | | | | | | |
| KRPS | (SEQ ID NO:4) | | | | | | |

**Figure 12 – Western Blots of Additional Antibodies**




1 GBS80 3 Arg-C

2 GBS80 F Arg-C

3 GBS80 F Asp-N

4 GBS80 F

# Figure 13 – Immunogenic Peptides in GBS 80

| | |
|---|---|
| GBS80 N-term | red |
| GBS80 C-term | blue |
| N-term mAbs on GBS80 | yellow |
| C-term mAbs on GBS80 | green + light blue |

| 1 | 11 | 21 | 31 | 41 | 51 | 61 | 71 |
|---|---|---|---|---|---|---|---|
| MAEVSQERPA | KTTVNIYKLQ | ADSYKSEITS | NGGIENKDGE | VISNYAKLGD | NVKGLQGVQF | KRYKVKTDIS | VDELKKLTTV |
| 81 | 91 | 101 | 111 | 121 | 131 | 141 | 151 |
| EAADAKVGTI | LEEGVSLPQK | TNAQGLVVDA | LDSKSNVRYL | YVEDLKNSPS | NITKAYAVPF | VLELPVANST | GTGFLSEINI |
| 161 | 171 | 181 | 191 | 201 | 211 | 221 | 231 |
| YPKNVVTDEP | KTDKDVKKLG | QDDAGYTIGE | EFKWFLKSTI | PANLGDYEKF | EITDKFADGL | TYKSVGKIKI | GSKTLNRDEH |
| 241 | 251 | 261 | 271 | 281 | 291 | 301 | 311 |
| YTIDEPTVDN | QNTLKITFKP | EKFKEIAELL | KGMTLVKNQD | ALDKATANTD | DAAFLEIPVA | STINEKAVLG | KAIENTFELQ |
| 321 | 331 | 341 | 351 | 361 | 371 | 381 | 391 |
| YDHTPDKADN | PKPSNPPRKP | EVHTGGKREV | KKDSTETQTL | GGAEFDLLAS | DGTAVKWTDA | LIKANTNKNY | IAGEAVTGQP |
| 401 | 411 | 421 | 431 | 441 | 451 | 461 | 471 |
| IKLKSHTDGT | FEIKGLAYAV | DANAEGTAVT | YKLKETKAPE | GYVIPDKEIE | FTVSQTSYNT | KPTDITVDSA | DATPDTIKNN |
| 481 | | | | | | | |
| KRPS | (SEQ ID NO:4) | | | | | | |

Figure 14 – Peptide MicroArray Layout

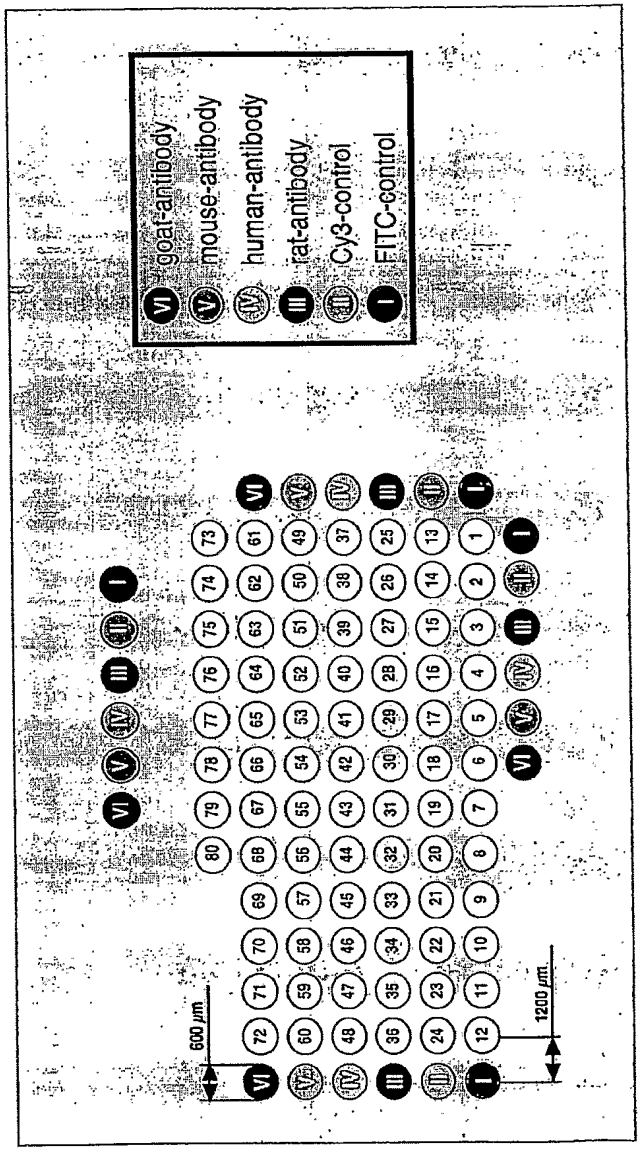
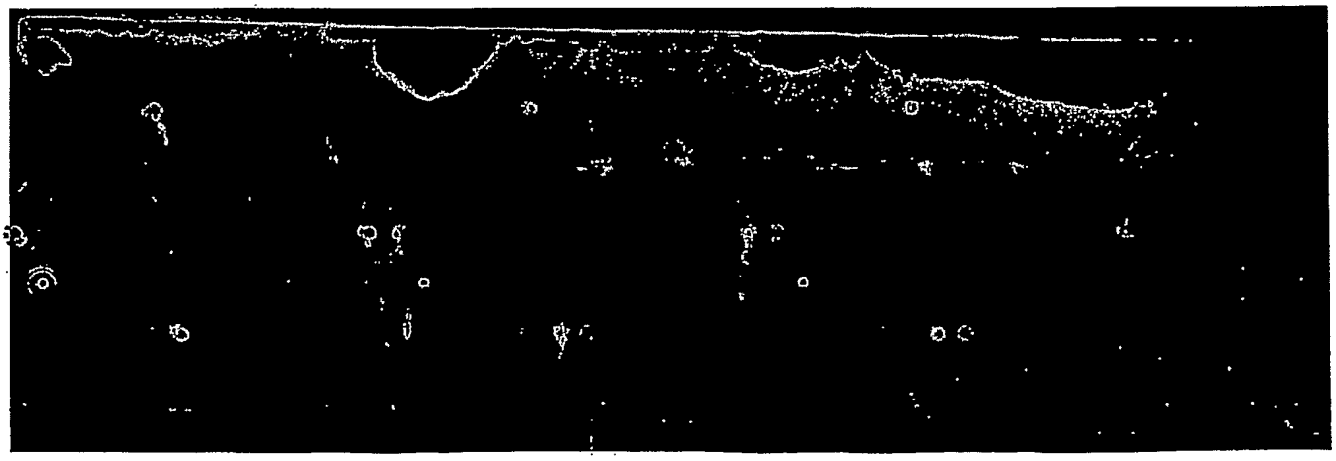


Figure 15 – Peptide MicroArray with Antibody

POSITIVE PEPTIDES

73) DALDSKSNVRYLY

75) SNVRYLYVEDLKN



# Figure 16 – Immunogenic Peptides in GBS 80

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 11 | 21 | 31 | 41 | 51 | 61 | 71 |
| MAEVSQERPA | KTTVNIYKLQ | ADSYKSEITS | NGGIENKDGE | VISNYAKLGD | NVKGLQGVQF | KRYKVKTDIS | VDELKKLTTV |
| 81 | 91 | 101 | 111 | 121 | 131 | 141 | 151 |
| EAADAKVGTI | LEEGVSLPQK | TNAQGLVVDA | EDSKSNVRYE | YVEDLKNSPS | NITKAYAVPF | VLELPVANST | GTGFLSEINI |
| 161 | 171 | 181 | 191 | 201 | 211 | 221 | 231 |
| YPKNVVTDEP | KTDKDVKKLG | QDDAGYTIGE | EFKWFLKSTI | PANLGDYEKF | EITDKFADGL | TYKSVGKIKI | GSKTLNRDEH |
| 241 | 251 | 261 | 271 | 281 | 291 | 301 | 311 |
| YTIDEPTVDN | QNTLKITFKP | EKFKEIAELL | KGMTLVKNQD | ALDKATANTD | DAAFLEIPVA | STINEKAVLG | KAIENTFELQ |
| 321 | 331 | 341 | 351 | 361 | 371 | 381 | 391 |
| YDHTPDKADN | PKPSNPPRKP | EVHTGGKRFV | KKDSTETQTL | GGAEFDLLAS | DGTAVKWTDA | LIKANTNKNY | IAGEAVTGQP |
| 401 | 411 | 421 | 431 | 441 | 451 | 461 | 471 |
| IKLKSHTDGT | FEIKGLAYAV | DANAEGTAVT | YKLKETKAPE | GYVIPDKEIE | FTVSQTSYNT | KPTDITVDSA | DATPDTIKNN |
| 481 | | | | | | | |
| KRPS | (SEQ ID NO:4) | | | | | | |